

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 00:45:53 ; Search time 231.102 Seconds  
(without alignments)  
2644.537 Million cell updates/sec

Title: US-09-647-780A-17  
Perfect score: 21  
Sequence: 1 catcatgcttttctctctggg 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 / Gapext 1.0

Searched: 2054640 seqs, 14551402878.residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pla.\*
- 35: em.htg.rod.\*
- 36: em.htg.man.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	21	100.0	21	6	AX014717	Sequence
C 2	19.4	92.4	2583	10	AF302075	AF302075 Mus muscu
C 3	19.4	92.4	2601	10	AF157106	AF157106 Mus muscu
C 4	19.4	92.4	2652	10	AF302076	AF302076 Mus muscu
C 5	19.4	92.4	2694	10	AF302077	AF302077 Mus muscu
C 6	19.4	92.4	2765	6	AX014701	AX014701 Sequence
C 7	19.4	92.4	2892	10	AF157105	AF157105 Mus muscu
C 8	19.4	92.4	2925	6	AX033272	AX033272 Sequence.
C 9	19.4	92.4	2925	10	AF176569	AF176569 Mus muscu
C 10	19.4	92.4	174953	2	AC094732	AC094732 Rattus no
C 11	19.4	92.4	182933	2	AC126977	AC126977 Rattus no
C 12	19.4	92.4	208249	2	AL607032	AL607032 Mus muscu
C 13	19.4	92.4	170363	2	AC112469	AC112469 Rattus no
C 14	19.4	92.4	203234	9	AL441943	AL441943 Human DNA
C 15	18.4	87.6	74600	9	AP000275	AP000275 Homo sapi
C 16	18.4	87.6	81877	8	T32B20	AF262041 Arabidops
C 17	18.4	87.6	100000	9	AP000037	AP000037 Homo sapi
C 18	18.4	87.6	100000	9	AP000105	AP000105 Homo sapi
C 19	18.4	87.6	100000	9	AP000181	AP000181 Homo sapi
C 20	18.4	87.6	162394	2	AC069518	AC069518 Homo sapi
C 21	18.4	87.6	183900	2	AC110789	AC110789 Homo sapi
C 22	18.4	87.6	191718	2	AC120471	AC120471 Rattus no
C 23	18.4	87.6	287820	2	AC096329	AC096329 Rattus no
C 24	18.4	87.6	340000	9	AP001714	AP001714 Homo sapi
C 25	18	85.7	143993	9	AC093607	AC093607 Homo sapi
C 26	18	85.7	159282	2	AC102933	AC102933 Mus muscu
C 27	18	85.7	166095	9	AC006052	AC006052 Homo sapi
C 28	18	85.7	174274	2	AC044905	AC044905 Homo sapi
C 29	17.8	84.8	2497	7	AH005854S4	M96254 Bacteriophe
C 30	17.8	84.8	4270	9	AK000787	AK000787 Homo sapi
C 31	17.8	84.8	111096	9	AC027304	AC027304 Homo sapi
C 32	17.8	84.8	119844	2	AC010044	AC010044 Drosophila
C 33	17.8	84.8	131823	2	AC010497	AC010497 Homo sapi
C 34	17.8	84.8	151160	2	AC120908	AC120908 Rattus no
C 35	17.8	84.8	153419	2	AC127038	AC127038 Rattus no
C 36	17.8	84.8	153926	9	AL160059	AL160059 Human DNA
C 37	17.8	84.8	154061	9	AC008524	AC008524 Homo sapi
C 38	17.8	84.8	174404	2	AC097028	AC097028 Rattus no
C 39	17.8	84.8	179647	2	AC108124	AC108124 Homo sapi
C 40	17.8	84.8	181323	9	AC108097	AC108097 Homo sapi
C 41	17.8	84.8	188120	2	AL392169	AL392169 Homo sapi
C 42	17.8	84.8	188207	2	AC010233	AC010233 Homo sapi
C 43	17.8	84.8	190700	2	AC128272	AC128272 Rattus no
C 44	17.8	84.8	191073	2	AC099760	AC099760 Mus muscu
C 45	17.8	84.8	191252	2	AC100345	AC100345 Mus muscu

ALIGNMENTS

RESULT 1  
AX014717  
LOCUS AX014717  
DEFINITION Sequence 17 from Patent.WO9953077.  
ACCESSION AX014717  
VERSION AX014717.1 GI:10040990  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM artificial construct.  
REFERENCE 1 (bases 1 to 21)  
AUTHORS Schwartz J.C., Gros C., Ouimet T., Rose C., Bonhomme M.C. and  
Facchinetti P.  
TITLE Novel nep ii membrane metalloprotease and its use for screening  
inhibitors useful in therapy

AX014717 21 bp DNA linear PAT 07-SEP-2000

JOURNAL Patent: WO 9953077-A 17 21-OCT-1999;  
 INST NAT SANTE RECH MED (FR); SCHWARTZ JEAN CHARLES (FR); GROS  
 CLAUDE (FR); OUMET TANJA (FR); ROSE CHRISTIANE (FR); BONHOMME  
 MARIE CHANTAL (FR); FACCHINETTI PATRICIA (FR)  
 FEATURES Location/Qualifiers  
 source  
 1..21  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="oligonucleotide"  
 2 a 6 c 4 g 9 t  
 BASE COUNT  
 ORIGIN  
 Query Match 100.0%; Score 21; DB 6; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CATCATGCTTTTCTCCTGGG 21  
 |||||  
 Db 1 CATCATGCTTTTCTCCTGGG 21  
 |||||  
 RESULT 2  
 AF302075/c 2583 bp mRNA linear ROD 11-JUN-2001  
 LOCUS  
 DEFINITION Mus musculus neprilysin-like peptidase alpha mRNA, complete cds.  
 ACCESSION AF302075  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 2583)  
 SHIROTANI K., TSUBUKI S., IWATA N., TAKAKI Y., HARIGAYA W.,  
 MARIYAMA K., KIRYU-SEO S., KIYAMA H., IWATA H., TOMITA T.,  
 Iwatsubo, T. and Saido, T.C.  
 Neprilysin degrades both amyloid beta peptides 1-40 and 1-42 most  
 rapidly and efficiently among thiorphan- and  
 phosphoramidon-sensitive endopeptidases  
 J. Biol. Chem. 276 (24), 21895-21901 (2001)  
 21293028  
 11278416  
 PUBLISHED  
 REFERENCE  
 2 (bases 1 to 2583)  
 SHIROTANI K. and SAIDO, T.C.  
 Direct Submission  
 TITLE  
 Submitted (02-SEP-2000) Proteolytic Neuroscience Laboratory, Brain  
 Science Institute, RIKEN, 2-1 Hirotsawa, Wako-shi, Saitama  
 351-0198, Japan  
 FEATURES  
 source  
 1..2583  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 25..2253  
 /note="endopeptidase"  
 /codon\_start=1  
 /product="neprilysin-like peptidase alpha"  
 /protein\_id="AA018446.1"  
 /db\_xref="GI:10505360"  
 /translation="MVERAGWCRRKSPGFVEYGLMVLILLGLAIVTLGVFYSIALRD  
 SSLKSDICTPSCVIAAARILENNQSRNPENFYQACGWLRRHVPETNSRYSVF  
 DILRDELVLKLVLEDSTQSRPAVEKATLRSCMNQSVIEKRDSPELVLKMWG  
 GWPVADKWNETMGLKWELEQLAVNSQFNRRVLDLFIWDDQNSRHVYIDQPT  
 LGMPREYFQEDNNHKVKAYLEPMTSVATMLRKDQNLKESAMVREMAEVLLET  
 HLANATVPEQRHDTALYHRMDLMELQERGLKGNWTLFIQNVLSSVEVLEFDEE  
 VVYGPYLENLEDDISYARTOMNYLVRLVLDRLIGLSQRFKEARVDYRKALYGT  
 TVEEVRRECYSVNSNMESAVGSLYIKRAFSDKSTVRELIEKIRSVFVDNLDELN  
 WMEDESKKAQKAMNREIQGPDYILEDDNNKHLDEEYSLTFYEDLYFENGLQNLK  
 NNAOSLKLKREKVDONLWIIIGAAVNVAFYSPNRNQIPEPAGILOPPFESKDQPSLN  
 FGGIGWIGHEITHGFDNGRNFKNQNLDMWNSFARHQFQQSCCMIIYQGNFSWE  
 LADNQNNGFISLGIENADNGGVQKAYKAYLRWLADGKDKDRLPGLNLTVAQLFFIN  
 AOVWCGSYRPEFAVQSIKTDVHSPKLYRVLGSLQNLPGFSEAFHCPGRSPMPMKRCR  
 IW"  
 665 a 667 c 736 g 515 t  
 BASE COUNT

ORIGIN  
 Query Match 92.4%; Score 19.4; DB 10; Length 2583;  
 Best Local Similarity 95.2%; Pred. No. 91;  
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CATCATGCTTTTCTCCTGGG 21  
 |||||  
 Db 895 CATCATGCTTTTCTCCTGGG 875  
 |||||  
 RESULT 3  
 AF157106/c 2601 bp mRNA linear ROD 25-NOV-1999  
 LOCUS  
 DEFINITION Mus musculus soluble-secreted endopeptidase delta mRNA,  
 alternatively spliced product, complete cds.  
 ACCESSION AF157106  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 2601)  
 IKEDA K., EMOTO N., RAHARJO S.B., NURHANTARI Y., SAIKI K.,  
 YOKOYAMA M. and MATSUO M.  
 Molecular identification and characterization of novel  
 membrane-bound metalloprotease, the soluble secreted form of which  
 hydrolyzes a variety of vasoactive peptides  
 J. Biol. Chem. 274 (45), 32469-32477 (1999)  
 20011457  
 10542292  
 PUBLISHED  
 REFERENCE  
 2 (bases 1 to 2601)  
 IKEDA K., EMOTO N. and MATSUO M.  
 Direct Submission  
 TITLE  
 Submitted (08-JUN-1999) International Center for Medical Research,  
 Kobe University School of Medicine, 7-5-1 Kusunoki, Chuo, Kobe  
 6500017, Japan  
 FEATURES  
 source  
 1..2601  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 59..2287  
 /note="SPP(delta); metalloprotease; alternatively spliced"  
 /codon\_start=1  
 /product="soluble secreted endopeptidase delta"  
 /protein\_id="AAF13153.1"  
 /db\_xref="GI:6467401"  
 /translation="MVERAGWCRRKSPGFVEYGLMVLILLGLAIVTLGVFYSIALRD  
 SSLKSDICTPSCVIAAARILENNQSRNPENFYQACGWLRRHVPETNSRYSVF  
 DILRDELVLKLVLEDSTQSRPAVEKATLRSCMNQSVIEKRDSPELVLKMWG  
 GWPVADKWNETMGLKWELEQLAVNSQFNRRVLDLFIWDDQNSRHVYIDQPT  
 LGMPREYFQEDNNHKVKAYLEPMTSVATMLRKDQNLKESAMVREMAEVLLET  
 HLANATVPEQRHDTALYHRMDLMELQERGLKGNWTLFIQNVLSSVEVLEFDEE  
 VVYGPYLENLEDDISYARTOMNYLVRLVLDRLIGLSQRFKEARVDYRKALYGT  
 TVEEVRRECYSVNSNMESAVGSLYIKRAFSDKSTVRELIEKIRSVFVDNLDELN  
 WMEDESKKAQKAMNREIQGPDYILEDDNNKHLDEEYSLTFYEDLYFENGLQNLK  
 NNAOSLKLKREKVDONLWIIIGAAVNVAFYSPNRNQIPEPAGILOPPFESKDQPSLN  
 FGGIGWIGHEITHGFDNGRNFKNQNLDMWNSFARHQFQQSCCMIIYQGNFSWE  
 LADNQNNGFISLGIENADNGGVQKAYKAYLRWLADGKDKDRLPGLNLTVAQLFFIN  
 AOVWCGSYRPEFAVQSIKTDVHSPKLYRVLGSLQNLPGFSEAFHCPGRSPMPMKRCR  
 IW"  
 655 a 681 c 748 g 517 t  
 BASE COUNT  
 ORIGIN  
 Query Match 92.4%; Score 19.4; DB 10; Length 2601;  
 Best Local Similarity 95.2%; Pred. No. 91;  
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CATCATGCTTTTCTCCTGGG 21  
 |||||  
 Db 929 CATCATGCTTTTCTCCTGGG 909  
 |||||

RESULT 4  
AF302076/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

2652 bp mRNA linear ROD 11-JUN-2001  
Mus musculus neprilysin-like peptidase beta mRNA, complete cds.

AF302076.1 GI:10505361  
Mus musculus  
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 2652)  
Shirotsani, K., Tsubuki, S., Iwata, N., Takaki, Y., Harigaya, W.,  
Maruyama, K., Kiryu-Seo, S., Kiyama, H., Iwata, H., Tomita, T.,  
Iwatsubo, T. and Saido, T.C.  
Neprilysin degrades both amyloid beta peptides 1-40 and 1-42 most  
rapidly and efficiently among thiorphan- and  
phosphoramidon-sensitive endopeptidases  
J. Biol. Chem. 276 (24), 21895-21901 (2001)  
21293028  
11278416  
PUBMED  
2 (bases 1 to 2652)  
Shirotsani, K. and Saido, T.C.  
Direct Submission  
TITLE  
JOURNAL

Neprilysin-like peptidase beta  
Submitted (02-SEP-2000) Proteolytic Neuroscience Laboratory, Brain  
Science Institute, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama  
351-0198, Japan

Location/Qualifiers  
1..2652  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
25..2322  
/note="endopeptidase"  
/codon\_start=1  
/product="neprilysin-like peptidase beta"  
/protein\_id="AAG18447.1"  
/db\_xref="GI:10505362"  
/translation="MVERAGWCRKKSPGVEYGLMVLVLLLLGAIIVTLGVFYSIGKQL  
PLLTSLHFWSDVTVKRALRDSLSKSDICTPSCVIAAARTLENDDSRNCFY  
QYACGGVLRHHVYPTNSRYSDVDFILRDEVLKGLVEDSTSOHPAYEKKATLYRS  
CMNQSVEIKRDESEPLSLVLMKVGWPMVDMKNETMGLKWELEQLAVLNSQFNRRVL  
IDLFTDNDQNSRHVYIDPTLGMPSREYFYQEDNNHKKRAYLEFMTSVATMLRK  
DONLSKESAMVEEAMAEVLELEHLANATVPQEKRDHDTALYHRMDLELOERFGLK  
FNWTFIONVLSVEVELFPDEEVVYVGYPIYLENLEIDIDSYSARTQNVLMVRLVD  
RIGLSQRFKEARVDYKALYGTVEEVRECVSVNSMESAGSLXIKKAFSKDS  
KSTVRELKIRSVFVNDLDELNMDEESKKRAQKAMIREQIGYDPDILENNKHL  
DEYSSTFLFYEDLYFENGQNLKNNKRAQSLKRLREKVDQNLITIGAAVNAFYSPARN  
QTVFFPAGLIQPPFQSKDQPSLNFQSGQCMYQYGNFSEWELADNQNVNGFTLGENIADNGGVQRAYKAYLRWLA  
FSAHFQOQSCMYYQYGNFSEWELADNQNVNGFTLGENIADNGGVQRAYKAYLRWLA  
DGKQDRLPLNLTVAOLFFINVAQVWCGSYRPEFAVQSIKTDVHSPKRYVLSGLQ  
LPGFEAFHCPGSPMHPMKRCRIW"

BASE COUNT 682 a 685 c 755 g 530 t  
ORIGIN

Query Match 92.4%; Score 19.4; DB 10; Length 2652;  
Best Local Similarity 95.2%; Pred. No. 91;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCTGGG 21  
||||| |||||||  
Db 964 CATCATGCTTTTCTCTCTGGG 944

RESULT 5  
AF302077/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

2694 bp mRNA linear ROD 11-JUN-2001  
Mus musculus neprilysin-like peptidase gamma mRNA, complete cds.

AF302077  
AF302077.1 GI:10505363  
Mus musculus  
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 2765)  
Schwartz, J.C., Gros, C., Oulmet, T., Rose, C., Bonhomme, M.C. and  
Faccinetti, P.  
Novel neprilysin membrane metalloprotease and its use for screening  
inhibitors useful in therapy  
Patent: WO 9953077-A 1 21-OCT-1999;  
INST NAT SANTE RECH MED (FR); SCHWARTZ JEAN CHARLES (FR); GROS

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 2694)  
Shirotsani, K., Tsubuki, S., Iwata, N., Takaki, Y., Harigaya, W.,  
Maruyama, K., Kiryu-Seo, S., Kiyama, H., Iwata, H., Tomita, T.,  
Iwatsubo, T. and Saido, T.C.  
Neprilysin degrades both amyloid beta peptides 1-40 and 1-42 most  
rapidly and efficiently among thiorphan- and  
phosphoramidon-sensitive endopeptidases  
J. Biol. Chem. 276 (24), 21895-21901 (2001)  
21293028  
11278416  
PUBMED  
2 (bases 1 to 2694)  
Shirotsani, K. and Saido, T.C.  
Direct Submission  
TITLE  
JOURNAL

Neprilysin-like peptidase gamma  
Submitted (02-SEP-2000) Proteolytic Neuroscience Laboratory, Brain  
Science Institute, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama  
351-0198, Japan

Location/Qualifiers  
1..2694  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
25..2364  
/note="endopeptidase"  
/codon\_start=1  
/product="neprilysin-like peptidase gamma"  
/protein\_id="AAG18448.1"  
/db\_xref="GI:10505364"  
/translation="MVERAGWCRKKSPGVEYGLMVLVLLLLGAIIVTLGVFYSIALRD  
SSLKSDICTPSCVIAAARTLENDDSRNCFYQYACGGVLRHHVYPTNSRYSVF  
DIIRDEVLKGLVEDSTSOHPAYEKKATLYRSCNQSVLEKRDSEPLSLVLMKVG  
GWPVDMKNETMGLKWELEQLAVLNSQFNRRVLIDLFTDNDQNSRHVYIDPT  
HLANATVPQEKRDHDTALYHRMDLELOERFGLKRDVSLCSPGCGTHSVDAQLELG  
LGMPSREYFYQEDNNHKKRAYLEFMTSVATMLRKQDLNLSKESAMVEEAMAEVLE  
NPPASDSRVGLKGFNWLFIQNVLSVEVELFPDEEVVYVGYPIYLENLEIDIDSYS  
RTMNVLMVRLVDRLGSLRSOREKARVDYKALYGTVEEVRECVSVNSMESAG  
VGSLEYIKRAFSDKSTVRELKIRSVFVNDLDELNMDEESKKRAQKAMIREQI  
GYDPDILENNKHLDEYSSTFLFYEDLYFENGQNLKNNKRAQSLKRLREKVDQNL  
GAAVNAFYSPARNQIVFPAGLIQPPFQSKDQPSLNFQSGQCMYQYGNFSEWELAD  
NQNNGFTLGENIADNGGVQRAYKAYLRWLA  
FSAHFQOQSCMYYQYGNFSEWELADNQNVNGFTLGENIADNGGVQRAYKAYLRWLA  
DGKQDRLPLNLTVAOLFFINVAQVWCGSYRPEFAVQSIKTDVHSPKRYVLSGLQ  
LPGFEAFHCPGSPMHPMKRCRIW"

BASE COUNT 586 a 700 c 766 g 542 t  
ORIGIN

Query Match 92.4%; Score 19.4; DB 10; Length 2694;  
Best Local Similarity 95.2%; Pred. No. 91;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCTGGG 21  
||||| |||||||  
Db 895 CATCATGCTTTTCTCTCTGGG 875

RESULT 6  
AX014701/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

2765 bp DNA linear PAT 07-SEP-2000  
Sequence 1 from Patent WO9953077.  
AX014701  
AX014701.1 GI:10040975  
black rat.  
Rattus rattus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 2765)  
Schwartz, J.C., Gros, C., Oulmet, T., Rose, C., Bonhomme, M.C. and  
Faccinetti, P.  
Novel neprilysin membrane metalloprotease and its use for screening  
inhibitors useful in therapy  
Patent: WO 9953077-A 1 21-OCT-1999;  
INST NAT SANTE RECH MED (FR); SCHWARTZ JEAN CHARLES (FR); GROS

```

CLAUDE (FR); OUTMET TANJA (FR); ROSE CHRISTIANE (FR); BONHOMME
MARIE CHANTAL (FR); FACHINETTI PATRICIA (FR)
Location/Qualifiers
1. 2765
/organism="Rattus rattus"
/db_xref="taxon:10117"
107..2431
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07576.1"
/db_xref="GI:10040976"
/translation="MGKSSGVMMERADNGRRRLGFVEGGLVLLTLLLMGATVTL
GVFYSIGKQLPNSLHSHVHEHTVVKVRLDRDSKSDICTTPTSCVIAAARILENMDOSRNPENFY
QSKKPCNFOFYACGGWLRHVIPETNSRYSDFDILRDELEVILKGVLEDSVQHRPA
VAKATLYRCMNOISVLEKRDSEPLNLVDMIGWPMADKWNMTGPKWELERQOLAV
LSQFNRRVLDLDFIWNDDQSSRHVIYIDPTLGMPSREYIFKEDSHRVREAYLOFM
TSVATMLRRDLNLPGEIDLVOEEMAOVLHLETHLANATVPOEKRDHVTALYHRMGLLE
LORFGLKGNWTLFIONVLSVVOVELLPNEEVVYGIPLYLENLEIIDVPAOTLON
YLVRLVLDIGLSLQRFKEARDYRKALYGTVMEEVWRRCVSVVNSMESAVGSLY
IKRAFSDKSLISSELKIRSVFVNDLNMDESKKKAQKAMNIRQIGIYDPI
ILEDNHRLDEEYSSLFSDLYFENGLQNNKNAQSLKLRKQVONLWIIIGAAVY
NATYSPNRNLVFPAGLQPPFFSKDQFQALNFGGIGMVGHEITHGFDNNGRNFEN
GNMLDMWNSFSAHFQOOSCMIVQYGNFWELADNONGVSTLGENIADNGVYRQA
YKAYLOWLAEGDRQRLPGLNLTVAOLFFINYAOVWCGSYRPEFAIQSIKTDVHSPK
YRVLSQNLPGSEAHCPGSPMHPMKRCRIW"
BASE COUNT 684 a 735 c 787 g 559 t
ORIGIN
Query Match 92.4%; Score 19.4; DB 6; Length 2765;
Best Local Similarity 95.2%; Pred. No. 90;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CATCATGCTTTTCTCTCTGGG 21
||||| |||||||
Db 1073 CATCATGCTTTTCTCTCTGGG 1053

RESULT 7
AF157105/c
LOCUS AF157105 2892 bp mRNA linear ROD 25-NOV-1999
DEFINITION Mus. musculus soluble secreted endopeptidase mRNA, complete cds.
ACCESSION AF157105
VERSION AF157105.1 GI:6467398
KEYWORDS
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 2892)
AUTHORS Ikeda, K., Emoto, N., Raharjo, S.B., Nurhantari, Y., Saiki, K.,
Yokoyama, M. and Matsuo, M.
TITLE Molecular identification and characterization of novel
membrane-bound metalloprotease, the soluble secreted form of which
hydrolyzes a variety of vasoactive peptides
J. Biol. Chem. 274 (45), 32469-32477 (1999)
JOURNAL
MEDLINE 20011457
PUBMED 10542292
REFERENCE 2 (bases 1 to 2892)
AUTHORS Ikeda, K., Emoto, N. and Matsuo, M.
TITLE Direct Submission
Submitted (08-JUN-1999) International Center for Medical Research,
Kobe University School of Medicine, 7-5-1 Kusunoki, Chuo, Kobe
6500017, Japan
FEATURES
source Location/Qualifiers
1. 2892
/organism="Mus musculus"
/db_xref="taxon:10090"
281..2578
/note="SEP: metalloprotease"
/codon_start=1
/product="soluble secreted endopeptidase"
/protein_id="AAFI3152.1"
/db_xref="GI:6467399"
CDS
Query Match 92.4%; Score 19.4; DB 6; Length 2925;
Best Local Similarity 95.2%; Pred. No. 90;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

/translation="MVERAGWCRKSPGVEYGLMVLVLLLLGAIIVTLGVFYSIGKQL
PLLSSLHFSNDEHTVVKRALRDSLSKSDICTTPTSCVIAAARILENMDOSRNPENFY
QYACGGWLRHVIPETNSRYSDFDILRDELEVILKGVLEDSVQHRPAVEKAKTLYRS
CMNOSVIEKRDSEPLLSVKMGWPMADKWNMTGMLKWELEQLAVLNSQFNRRVL
IDLFINDDQSSRHVIYIDPTLGMPSREYIFQEDNNHKVRKAYLEFMTSVATMLRK
DONLSKESAMVREEMAELFLETHLANATVPOEKRDHVTALYHRMDLMELQERFGLKG
FNWTLFIONVLSVVEVLPDEEVVYGIPLYLENLEIIDSYSAFTMONYLVRLVLD
RIGLSLQRFKEARDYRKALYGTVMEEVWRRCVSVVNSMESAVGSLYIKRAFSDKS
KSTVKLEKIRSVFVNDLNMDESKKKAQKAMNIRQIGIYDPILEDNNKHL
DEEYSSLFYEDLYFENGLQNNKNAQSLKLRKQVONLWIIIGAAVYNAFYSPNRN
QIVFPAGLQPPFFSKDQFQALNFGGIGMVGHEITHGFDNNGRNFENGLDMWNS
FSARHFOOSCMIVQYGNFWELADNONGVSTLGENIADNGVYRQAYKAYLRWLA
DGGKDRLPGLNLTVAOLFFINYAOVWCGSYRPEFAVQSIKTDVHSPKLYRVLSGLQ
LPGSEAHCPGSPMHPMKRCRIW"
BASE COUNT 717 a 774 c 827 g 574 t
ORIGIN
Query Match 92.4%; Score 19.4; DB 10; Length 2892;
Best Local Similarity 95.2%; Pred. No. 90;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CATCATGCTTTTCTCTCTGGG 21
||||| |||||||
Db 1220 CATCATGCTTTTCTCTCTGGG 1200

RESULT 8
AX033272/c
LOCUS AX033272 2925 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 12 from Patent WO0047750.
ACCESSION AX033272
VERSION AX033272.1 GI:10280087
KEYWORDS house mouse;
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 2925)
AUTHORS Boileau, G. and Desgroseillers, L.
TITLE New metalloproteases of the neprilysin family
Patent: WO 0047750-A.12 17-AUG-2000;
JOURNAL BOILEAU GUY (CA) ; DESGROSEILLERS LUC (CA) ; UNIVERSITE DE MONTREAL
(CA)
FEATURES
source Location/Qualifiers
1. 2925
/organism="Mus musculus"
/db_xref="taxon:10090"
332..2629
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09977.1"
/db_xref="GI:10280088"
/translation="MVERAGWCRKSPGVEYGLMVLVLLLLGAIIVTLGVFYSIGKQL
PLLSSLHFSNDEHTVVKRALRDSLSKSDICTTPTSCVIAAARILENMDOSRNPENFY
QYACGGWLRHVIPETNSRYSDFDILRDELEVILKGVLEDSVQHRPAVEKAKTLYRS
CMNOSVIEKRDSEPLLSVKMGWPMADKWNMTGMLKWELEQLAVLNSQFNRRVL
IDLFINDDQSSRHVIYIDPTLGMPSREYIFQEDNNHKVRKAYLEFMTSVATMLRK
DONLSKESAMVREEMAELFLETHLANATVPOEKRDHVTALYHRMDLMELQERFGLKG
FNWTLFIONVLSVVEVLPDEEVVYGIPLYLENLEIIDSYSAFTMONYLVRLVLD
RIGLSLQRFKEARDYRKALYGTVMEEVWRRCVSVVNSMESAVGSLYIKRAFSDKS
KSTVKLEKIRSVFVNDLNMDESKKKAQKAMNIRQIGIYDPILEDNNKHL
DEEYSSLFYEDLYFENGLQNNKNAQSLKLRKQVONLWIIIGAAVYNAFYSPNRN
QIVFPAGLQPPFFSKDQFQALNFGGIGMVGHEITHGFDNNGRNFENGLDMWNS
FSARHFOOSCMIVQYGNFWELADNONGVSTLGENIADNGVYRQAYKAYLRWLA
DGGKDRLPGLNLTVAOLFFINYAOVWCGSYRPEFAVQSIKTDVHSPKLYRVLSGLQ
LPGSEAHCPGSPMHPMKRCRIW"
BASE COUNT 710 a 797 c 836 g 582 t
ORIGIN
Query Match 92.4%; Score 19.4; DB 6; Length 2925;
Best Local Similarity 95.2%; Pred. No. 90;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```



QY 1 CATCATGCTTTTCTCTCTGGG 21  
 ||||| ||||| ||||| ||||| |||||  
 Db 1271 CATCATGCTTTTCTCTCTGGG 1251

## RESULT 9

AF176569/c

LOCUS

DEFINITION

AF176569 2925 bp mRNA linear ROD 11-MAY-2000  
 Mus musculus neprilysin-like metalloproteinase 1 (N11) mRNA,  
 complete cds.

ACCESSION

AF176569

VERSION

AF176569.1 GI:7769082

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

20; Conservative

0; Mismatches

1; Indels

0; Gaps

0;

QY 1 CATCATGCTTTTCTCTCTGGG 21

Db 1271 CATCATGCTTTTCTCTCTGGG 1251

RESULT 10

AC094732

LOCUS

DEFINITION

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus

1 (bases 1 to 2925)

Chadgar,G., Ruchon,A.F., Carpentier,M., Marcinkiewicz,M.,

Seidah,N.G., Crine,P., Desgroselliers,L. and Boileau,G.

Molecular cloning and biochemical characterization of a new mouse

testis soluble-zinc-metalloproteinase of the neprilysin family

Biochem. J. 347 (Pt 2), 419-429 (2000)

21060448

10749671

2 (bases 1 to 2925)

Chadgar,G., Ruchon,A.F., Desgroselliers,L. and Boileau,G.

Submitted (09-AUG-1999) Biochemistry, University of Montreal, 2900

Edouard Montpetit, Montreal, Que H3T 1J4, Canada

Location/Qualifiers

1. .2925

/organism="Mus musculus"

/db\_xref="taxon:10090"

/tissue\_type="testis"

1. .2925

/gene="N11"

332. .2629

/genes="N11"

/note="N11; metallo-endopeptidase"

/codon\_start=1

/product="neprilysin-like metalloproteinase 1"

/protein\_id="AAF69247.1"

/db\_xref="GI:7769083"

/translation="MVERAGCCKRKKSPGVEYGLMVLILLLLLLGAVTLGVFYSIGKOL  
 PLITSLHPSWDETVVKRALRDSLSKIDICTTPSCVIAAARILENDQSRNCFNY  
 QYACGSLRHVTPENRSYSPEDILRDEVLKGLVLEDSQHPAPVSKAKTLVES  
 CMQSVIERKDSPLSLVKWCGWPFVAMDKNETMGLKWELEQLAVLSQFNRRVL  
 DQNLKSAWREMAEVLLETHLANATVPQKRHDVTALYHRMDLMEQERFGLK  
 FNWTLFQNLVSVLEVPDEEVVYGIPILENLEIDISYARTMONTLYLVRLVLD  
 RIGLSQREKARDYRKALYGTVEVRWRECVSYVNSMESAVGSLYIKRAFSDS  
 KTVRELIRKISVFDNLDLANNDEESKKKAQKAMNTREQIGVDPDYILEDNKHL  
 DEYSLTTFEDLYFENGLONLKNNAQSRSLKLRKVDQNLWIIIGAAVNAFSPNRN  
 QIVFPAGIOLPPFSKDPQSLNPGFGIGVIGHEITGFDNDGNGFNKGNMLDWSN  
 FSARHPQSQSCMIYQYGNFSLWELANQYNGFTLGTENADNGVQYAKYLRLWA  
 DGKDKRLPLNLNTYALQFFINTAQVWCGSRPEFAVQSIKTDVHSPKLYRVLGSLON  
 LPGSFARHCPRGSPHMPMKRCRIW"

BASE COUNT 710 a 797 c 836 g 582 t

ORIGIN

Query Match 92.4%; Score 19.4; DB 10; Length 2925;

Best Local Similarity 95.2%; Pred. No. 90;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCTGGG 21

Db 1271 CATCATGCTTTTCTCTCTGGG 1251

RESULT 10

AC094732

LOCUS

DEFINITION

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus

1 (bases 1 to 174953)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Alsbrooks,S.I., Amarantunge,H.C., Are,J.R., Banks,T., Barbara,J.,

Benton,J., Bimege,K., Blankenburg,K., Bonnin,D., Bouck,J.,

Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,

Burckett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,

Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,

Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,

Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,

Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,

Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,

Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,

Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,

Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,

Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,

Hamilton,K., Harris,C., Harris,K., Hart,M., Havliak,P., Hawes,A.,

Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,

Hollins,B., Homs,J.F., Howard,S., Huber,J., Hulyk,S., Hume,J.,

Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,

Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,

Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,

Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,

Loulseged,H., Lozada,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,

Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,

Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,

Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,

Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,

Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,

Ogih,M., Okwunodu,G., Oragunye,N., Owiedo,R., Pace,A., Payton,B.,

Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,

Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojuben,I., Rolfe,M.,

Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,

Sisson,I., Sodergren,E., Sonaik,R., Sparks,A., Stanley,H.,

Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,

Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,

Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,

Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,

Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,

Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,

Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 174953)

Worley,K.C.

Direct Submission

Submitted (15-SEP-2001) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Dec 20, 2001 this sequence version replaced gi:15624568.

Center: Genome Center

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GHG

Center clone name: CH230-516

Summary Statistics

Assembly program: Phrap; version 0.990329First call to

findPhrapList

Consensus quality: 152255 bases at least Q40

Consensus quality: 158448 bases at least Q30

Consensus quality: 164461 bases at least Q20

DEFINITION

Rattus norvegicus clone CH230-516, \*\*\* SEQUENCING IN PROGRESS \*\*\*

56 unordered pieces.

AC094732

VERSION

GI:17941511

KEYWORDS

HTG; HTGS\_PHASE1

SOURCE

Rattus norvegicus

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

BASE COUNT

ORIGIN

Query Match 92.4%; Score 19.4; DB 10; Length 2925;

Best Local Similarity 95.2%; Pred. No. 90;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCTGGG 21

Db 1271 CATCATGCTTTTCTCTCTGGG 1251

RESULT 10

AC094732

LOCUS

DEFINITION

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus

1 (bases 1 to 174953)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Alsbrooks,S.I., Amarantunge,H.C., Are,J.R., Banks,T., Barbara,J.,

Benton,J., Bimege,K., Blankenburg,K., Bonnin,D., Bouck,J.,

Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,

Burckett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,

Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,

Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,

Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,

Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,

Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,

Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,

Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,

Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,

Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,

Hamilton,K., Harris,C., Harris,K., Hart,M., Havliak,P., Hawes,A.,

Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,

Hollins,B., Homs,J.F., Howard,S., Huber,J., Hulyk,S., Hume,J.,

Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,

Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,

Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,

Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,

Loulseged,H., Lozada,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,

Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,

Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,

Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,

Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,

Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,

Ogih,M., Okwunodu,G., Oragunye,N., Owiedo,R., Pace,A., Payton,B.,

Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,

Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojuben,I., Rolfe,M.,

Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,

Sisson,I., Sodergren,E., Sonaik,R., Sparks,A., Stanley,H.,

Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,

Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,

Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,

Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,

Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,

Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,

Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 174953)

Worley,K.C.

Direct Submission

Submitted (15-SEP-2001) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Dec 20, 2001 this sequence version replaced gi:15624568.

Center: Genome Center

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GHG

Center clone name: CH230-516

Summary Statistics

Assembly program: Phrap; version 0.990329First call to

findPhrapList

Consensus quality: 152255 bases at least Q40

Consensus quality: 158448 bases at least Q30

Consensus quality: 164461 bases at least Q20

Estimated insert size: 155965; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-gel estimation  
 Quality coverage: 2.2x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 NOTE: this is a 'working draft' sequence. It currently  
 consists of 56 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1 13782: contig of 13782 bp in length  
 13783 13882: contig of unknown length  
 13883 23287: contig of 9405 bp in length  
 23287 23387: gap of unknown length  
 23387 23888: contig of 4694 bp in length  
 23888 28081: contig of 4694 bp in length  
 28081 28181: gap of unknown length  
 28181 33807: contig of 5626 bp in length  
 33807 33907: gap of unknown length  
 33907 39271: contig of 5364 bp in length  
 39271 39371: gap of unknown length  
 39371 44270: contig of 4899 bp in length  
 44270 44370: gap of unknown length  
 44370 47223: contig of 3353 bp in length  
 47223 47724: gap of unknown length  
 47724 47824: contig of 5604 bp in length  
 47824 53528: gap of unknown length  
 53528 57303: contig of 3776 bp in length  
 57303 57403: gap of unknown length  
 57403 62018: contig of 4615 bp in length  
 62018 62118: gap of unknown length  
 62118 66776: contig of 4558 bp in length  
 66776 70201: contig of 3425 bp in length  
 70201 70301: gap of unknown length  
 70301 74082: contig of 3781 bp in length  
 74082 74182: gap of unknown length  
 74182 78306: contig of 4124 bp in length  
 78306 78406: gap of unknown length  
 78406 81222: contig of 2816 bp in length  
 81222 81322: gap of unknown length  
 81322 84552: contig of 3230 bp in length  
 84552 84652: gap of unknown length  
 84652 88839: contig of 4187 bp in length  
 88839 88939: gap of unknown length  
 88939 92256: contig of 3317 bp in length  
 92256 92356: gap of unknown length  
 92356 94995: contig of 2639 bp in length  
 94995 95095: gap of unknown length  
 95095 98465: contig of 3370 bp in length  
 98465 98565: gap of unknown length  
 98565 101632: contig of 3067 bp in length  
 101632 101732: gap of unknown length  
 101732 104939: contig of 3207 bp in length  
 104939 105039: gap of unknown length  
 105039 109248: contig of 4209 bp in length  
 109248 109348: gap of unknown length  
 109348 112435: contig of 3087 bp in length  
 112435 112535: gap of unknown length  
 112535 114676: contig of 2141 bp in length  
 114676 114776: gap of unknown length  
 114776 117567: contig of 2791 bp in length  
 117567 117667: gap of unknown length  
 117667 119209: contig of 1542 bp in length  
 119209 119309: gap of unknown length  
 119309 121797: contig of 2488 bp in length  
 121797 121897: gap of unknown length  
 121897 124293: contig of 2396 bp in length  
 124293 124394: gap of unknown length  
 124394 126428: contig of 2035 bp in length

126429 126528: gap of unknown length  
 126528 129525: contig of 2997 bp in length  
 129525 129625: gap of unknown length  
 129625 132334: contig of 2709 bp in length  
 132334 132434: gap of unknown length  
 132434 135274: contig of 2840 bp in length  
 135274 135374: gap of unknown length  
 135374 138074: contig of 2700 bp in length  
 138074 138174: gap of unknown length  
 138174 139985: contig of 1811 bp in length  
 139985 140085: gap of unknown length  
 140085 142273: contig of 2188 bp in length  
 142273 142373: gap of unknown length  
 142373 143598: contig of 1225 bp in length  
 143598 143698: gap of unknown length  
 143698 145434: contig of 1736 bp in length  
 145434 145534: gap of unknown length  
 145534 146985: contig of 1451 bp in length  
 146985 147085: gap of unknown length  
 147085 148099: contig of 1014 bp in length  
 148099 148199: gap of unknown length  
 148199 150315: contig of 2716 bp in length  
 150315 151015: gap of unknown length  
 151015 152501: contig of 1486 bp in length  
 152501 152601: gap of unknown length  
 152601 154010: contig of 1409 bp in length  
 154010 154110: gap of unknown length  
 154110 155758: contig of 1648 bp in length  
 155758 155858: gap of unknown length  
 155858 157622: contig of 1764 bp in length  
 157622 157722: gap of unknown length  
 157722 159428: contig of 1706 bp in length  
 159428 159528: gap of unknown length  
 159528 161209: contig of 1681 bp in length  
 161209 161309: gap of unknown length  
 161309 163413: contig of 2104 bp in length  
 163413 163513: gap of unknown length  
 163513 164702: contig of 1189 bp in length  
 164702 164802: gap of unknown length  
 164802 165998: contig of 1196 bp in length  
 165998 166098: gap of unknown length  
 166098 167412: contig of 1314 bp in length  
 167412 167512: gap of unknown length  
 167512 169231: contig of 1719 bp in length  
 169231 169331: gap of unknown length  
 169331 170334: contig of 1203 bp in length  
 170334 170535: gap of unknown length  
 170535 172047: contig of 1413 bp in length  
 172047 172147: gap of unknown length  
 172147 173509: contig of 1362 bp in length  
 173509 173609: gap of unknown length  
 173609 174953: contig of 1344 bp in length.  
 174953 173610: Location/Qualifiers

## FEATURES

Query Match 92.4%; Score 19.4; DB 2; Length 174953;  
 Best Local Similarity 95.2%; Pred. No. 68;  
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCCTGGG 21  
 Db 58747 CATCATGCTTTTCTCCTGGG 58767

RESULT 11  
 AC126977

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AC126977 182933 bp DNA linear HTG 17-JUL-2002  
 Rattus norvegicus clone CH230-11A4, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
 78 unordered pieces.  
 AC126977.2 GI:21745593  
 HTG: HTGS\_PHASE1.  
 Norway rat.  
 Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

# REFERENCE AUTHORS

1 (bases 1 to 182933)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
Barbaria,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,  
Bouch,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Haves,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissegh,H.,  
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokkwo,S., Ogih,M., Okwuonu,G.,  
Oraguaye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojupokan,I., Rolfe,M., Ruiz,S., Savory,G.,  
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,  
Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczyk,R., Woodson,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G., and Gibbs,R.

# TITLE JOURNAL

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Submitted (11-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 182933)

# REFERENCE AUTHORS TITLE JOURNAL

Submitted (17-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 14, 2002 this sequence version replaced gi:21731393

# COMMENT

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GKAG  
Center clone name: CH230-11A4  
----- Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 113910 bases at least Q40  
Consensus quality: 119514 bases at least Q30

Consensus quality: 124341 bases at least Q20

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
NOTE: This is a 'working draft' sequence. It currently  
consists of 78 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 1024: contig of 1024 bp in length  
1025 1124: gap of unknown length  
1125 1124: contig of 1020 bp in length  
2145 2244: gap of unknown length  
2245 3928: contig of 1684 bp in length  
3929 4028: gap of unknown length  
4029 5293: contig of 1265 bp in length  
5294 5393: gap of unknown length  
5394 7069: contig of 1676 bp in length  
7070 7169: gap of unknown length  
7170 8762: contig of 1593 bp in length  
8763 8862: gap of unknown length  
8863 10240: contig of 1378 bp in length  
10241 10340: gap of unknown length  
10341 11621: contig of 1281 bp in length  
11622 11721: gap of unknown length  
11722 13291: contig of 1570 bp in length  
13292 13391: gap of unknown length  
13392 14660: contig of 1269 bp in length  
14661 14760: gap of unknown length  
14761 15851: contig of 1091 bp in length  
15852 17000: contig of 1049 bp in length  
17001 17100: gap of unknown length  
17101 18132: contig of 1032 bp in length  
18133 18232: gap of unknown length  
18233 19666: contig of 1434 bp in length  
19667 19766: gap of unknown length  
19767 20845: contig of 1079 bp in length  
20846 20945: gap of unknown length  
20946 22317: contig of 1372 bp in length  
22318 22417: gap of unknown length  
22418 24395: contig of 1978 bp in length  
24396 24495: gap of unknown length  
24496 25723: contig of 1228 bp in length  
25724 25823: gap of unknown length  
25824 27450: contig of 1627 bp in length  
27451 27550: gap of unknown length  
27551 28762: contig of 1212 bp in length  
28763 28862: gap of unknown length  
28863 30007: contig of 1145 bp in length  
30008 30107: gap of unknown length  
30108 31349: contig of 1242 bp in length  
31350 31449: gap of unknown length  
31450 32794: contig of 1345 bp in length  
32795 32894: gap of unknown length  
32895 34562: contig of 1668 bp in length  
34563 34662: gap of unknown length  
34663 35689: contig of 1027 bp in length  
35690 35789: gap of unknown length  
35790 37083: contig of 1294 bp in length  
37084 37183: gap of unknown length  
37184 38403: contig of 1220 bp in length  
38404 38503: gap of unknown length  
38504 39564: contig of 1061 bp in length  
39565 39664: gap of unknown length  
39665 40851: contig of 1187 bp in length  
40852 40951: gap of unknown length  
40952 42028: contig of 1077 bp in length  
42029 42128: gap of unknown length  
42129 44335: contig of 2207 bp in length

44336 gap of unknown length  
 44336 contig of 1833 bp in length  
 46268 gap of unknown length  
 46268 contig of 1521 bp in length  
 47889 gap of unknown length  
 47889 contig of 1464 bp in length  
 49453 gap of unknown length  
 49453 contig of 1294 bp in length  
 50847 gap of unknown length  
 50847 contig of 1982 bp in length  
 52929 gap of unknown length  
 52929 contig of 2052 bp in length  
 53030 gap of unknown length  
 53030 contig of 1978 bp in length  
 55182 gap of unknown length  
 55182 contig of 1423 bp in length  
 57160 gap of unknown length  
 57160 contig of 1581 bp in length  
 58682 gap of unknown length  
 58682 contig of 2198 bp in length  
 60364 gap of unknown length  
 60364 contig of 1481 bp in length  
 62662 gap of unknown length  
 62662 contig of 1706 bp in length  
 64243 gap of unknown length  
 64243 contig of 1350 bp in length  
 66049 gap of unknown length  
 66049 contig of 1886 bp in length  
 67499 gap of unknown length  
 67499 contig of 1861 bp in length  
 69485 gap of unknown length  
 69485 contig of 1352 bp in length  
 71446 gap of unknown length  
 71446 contig of 2544 bp in length  
 72998 gap of unknown length  
 72998 contig of 1665 bp in length  
 75542 gap of unknown length  
 75542 contig of 1840 bp in length  
 77307 gap of unknown length  
 77307 contig of 2204 bp in length  
 79247 gap of unknown length  
 79247 contig of 2589 bp in length  
 81651 gap of unknown length  
 81651 contig of 1403 bp in length  
 84239 gap of unknown length  
 84239 contig of 1392 bp in length  
 85743 gap of unknown length  
 85743 contig of 1392 bp in length

Query Match 92.4%; Score 19.4; DB 2; Length 182933;  
 Best Local Similarity 95.2%; Pred. No. 68;  
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 CATCATGCTTTTCTCTCTGGG 21  
 Db 7194 CATCATGCTTTTCTCTCTGGG 7214  
 RESULT 12  
 AL607032/c  
 LOCUS 208249 bp DNA linear HTG 24-JUL-2002  
 DEFINITION Mus musculus chromosome 4 clone RP23-15L19, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\*, 2 unordered pieces.  
 ACCESSION AL607032  
 VERSION AL607032.15 GI:21955491  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_ACTIVEFIN; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus  
 1 (bases 1 to 208249)  
 Direct Submission  
 Tracey, A.  
 Submitted (23-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

## COMMENT

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On Jul 25, 2002 this sequence version replaced gi:21627906.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquery@sanger.ac.uk  
 ----- Project Information  
 Center project name: BM15L19  
 ----- Summary Statistics  
 Assembly program: XGAP4; version 4.5  
 Sequencing vector: plasmid; L08752; 100% of reads  
 Chemistry: Dye-terminator; 11% of reads  
 Consensus quality: 207834 bases at least Q40  
 Consensus quality: 207956 bases at least Q30  
 Consensus quality: 208060 bases at least Q20  
 Insert size: 219051; 2.3% error; agarose-fp  
 Quality coverage: 12.41x in Q20 bases; sum-of-contigs Quality  
 coverage: 12.08x in Q20 bases; agarose-fp

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 2 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 137303: contig of 137303 bp in length  
 \* 137304 137403: gap of 100 bp  
 \* 137404 208249: contig of 70846 bp in length.

## FEATURES

source  
 1. 208249  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="4"  
 /clone="RP23-15L19"  
 /clone\_lib="RPCI-23"  
 1. 137303  
 /note="assembly\_fragment:04071"  
 137404. 208249  
 /note="assembly\_fragment:01838"  
 clone\_end:T7  
 vector\_side:right  
 BASE COUNT 54098 a 53647 c 51328 g 49076 t 100 others  
 ORIGIN

Query Match 92.4%; Score 19.4; DB 2; Length 208249;  
 Best Local Similarity 95.2%; Pred. No. 67;  
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CATCATGCTTTTCTCTCTGGG 21

Db 175900 CATCATGCTTTTCTCTCTGGG 175880

## RESULT 13

AC112469  
 LOCUS 170363 bp DNA linear HTG 17-JUL-2002  
 DEFINITION Rattus norvegicus clone.CH230-118C9, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\* 75 unordered pieces.  
 ACCESSION AC112469  
 VERSION AC112469.3 GI:21745528  
 KEYWORDS HTG; HTGS\_PHASE1.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 170363)  
 REFERENCES  
 AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,  
 Barbaria, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D.,  
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,  
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,  
 Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jollivet, S., Joudah, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,  
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G.,  
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,  
 Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,  
 Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansley, J., Taylor, C., Taylor, T., Telifrod, B., Thomas, N., Thomas, S.,  
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
 Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.

TITLE  
JOURNALREFERENCE  
Unpublished

2 (bases 1 to 170363)

TITLE  
JOURNALREFERENCE  
Unpublished

2 (bases 1 to 170363)

REFERENCE  
AUTHORSSubmitted (21-FEB-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USATITLE  
JOURNAL

3 (bases 1 to 170363)

REFERENCE  
AUTHORSSubmitted (17-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

## COMMENT

On Jul 14, 2002 this sequence version replaced gi:20303357.  
 -----  
 Center: Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 -----  
 Project Information  
 -----  
 Center project name: GSKJ  
 Center clone name: CH230-118C9  
 -----  
 Summary Statistics  
 -----  
 Sequencing vector: Plasmid;  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 104743 bases at least Q40  
 Consensus quality: 111029 bases at least Q30  
 Consensus quality: 116478 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))  
 \* NOTE: This is a 'working draft' sequence. It currently

\* consists of 75 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence.  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1041: contig of 1041 bp in length  
 \* 1042: gap of unknown length  
 \* 1141: contig of 1075 bp in length  
 \* 1142: gap of unknown length  
 \* 2217: contig of 1175 bp in length  
 \* 3491: gap of unknown length  
 \* 3591: gap of unknown length  
 \* 4801: contig of 1210 bp in length  
 \* 4901: gap of unknown length  
 \* 4902: contig of 1304 bp in length  
 \* 6305: gap of unknown length  
 \* 6306: contig of 1531 bp in length  
 \* 7837: gap of unknown length  
 \* 9303: contig of 1367 bp in length  
 \* 9304: gap of unknown length  
 \* 11382: contig of 1979 bp in length  
 \* 11482: gap of unknown length  
 \* 12552: contig of 1070 bp in length  
 \* 12652: gap of unknown length  
 \* 14154: contig of 1502 bp in length  
 \* 14254: gap of unknown length  
 \* 14259: contig of 1043 bp in length  
 \* 15298: gap of unknown length  
 \* 15397: contig of 1934 bp in length  
 \* 17331: gap of unknown length  
 \* 17332: contig of 1399 bp in length  
 \* 18830: gap of unknown length  
 \* 18831: contig of 1499 bp in length  
 \* 20429: gap of unknown length  
 \* 20529: contig of 1100 bp in length  
 \* 20530: gap of unknown length  
 \* 21630: contig of 1317 bp in length  
 \* 21729: gap of unknown length  
 \* 23046: contig of 1317 bp in length  
 \* 23047: gap of unknown length  
 \* 23147: contig of 1729 bp in length  
 \* 24875: gap of unknown length  
 \* 24975: contig of 1596 bp in length  
 \* 26571: gap of unknown length  
 \* 26572: contig of 1424 bp in length  
 \* 26572: gap of unknown length  
 \* 28095: contig of 1553 bp in length  
 \* 28195: gap of unknown length  
 \* 28196: contig of 1553 bp in length  
 \* 29748: gap of unknown length  
 \* 29848: contig of 2267 bp in length  
 \* 32115: gap of unknown length  
 \* 32215: contig of 1672 bp in length  
 \* 33066: contig of 1291 bp in length  
 \* 33507: gap of unknown length  
 \* 35613: contig of 2007 bp in length  
 \* 35713: gap of unknown length  
 \* 37385: contig of 1672 bp in length  
 \* 37386: gap of unknown length  
 \* 37485: contig of 1082 bp in length  
 \* 38567: gap of unknown length  
 \* 38668: contig of 1297 bp in length  
 \* 39965: gap of unknown length  
 \* 40064: contig of 1424 bp in length  
 \* 41488: gap of unknown length  
 \* 41588: contig of 2150 bp in length  
 \* 43738: gap of unknown length  
 \* 43739: contig of 1170 bp in length  
 \* 45008: gap of unknown length  
 \* 45108: contig of 1130 bp in length  
 \* 46238: gap of unknown length  
 \* 46338: contig of 2211 bp in length  
 \* 48549: gap of unknown length  
 \* 48649: contig of 1549 bp in length  
 \* 48550: gap of unknown length  
 \* 50198: contig of 1549 bp in length  
 \* 50298: gap of unknown length  
 \* 51619: contig of 1321 bp in length  
 \* 51620: gap of unknown length

\* 51720 52948: contig of 1229 bp in length  
 \* 52949 53048: gap of unknown length  
 \* 53049 55060: contig of 2012 bp in length  
 \* 55061 55160: gap of unknown length  
 \* 55161 56970: contig of 1810 bp in length  
 \* 56971 57070: gap of unknown length  
 \* 57071 58683: contig of 1613 bp in length  
 \* 58684 58783: gap of unknown length  
 \* 58784 60555: contig of 1772 bp in length  
 \* 60556 60655: gap of unknown length  
 \* 60656 62556: contig of 1901 bp in length  
 \* 62557 62656: gap of unknown length  
 \* 62657 65132: contig of 2476 bp in length  
 \* 65133 65232: gap of unknown length  
 \* 65233 66689: contig of 1457 bp in length  
 \* 66690 66789: gap of unknown length  
 \* 66790 68045: contig of 1256 bp in length  
 \* 68046 68145: gap of unknown length  
 \* 68146 69483: contig of 1338 bp in length  
 \* 69484 69583: gap of unknown length  
 \* 69584 72087: contig of 2504 bp in length  
 \* 72088 72187: gap of unknown length  
 \* 72188 73395: contig of 1208 bp in length  
 \* 73396 73495: gap of unknown length  
 \* 73496 74977: contig of 1482 bp in length  
 \* 74978 75077: gap of unknown length  
 \* 75078 76750: contig of 1673 bp in length  
 \* 76751 76850: gap of unknown length  
 \* 76851 79029: contig of 2179 bp in length  
 \* 79030 79129: gap of unknown length  
 \* 79130 81139: contig of 2010 bp in length  
 \* 81140 81239: gap of unknown length  
 \* 81240 83895: contig of 2656 bp in length  
 \* 83896 83995: gap of unknown length  
 \* 83996 85591: contig of 1596 bp in length  
 \* 85592 85691: gap of unknown length  
 \* 85692 87734: contig of 2043 bp in length  
 \* 87735 87834: gap of unknown length  
 \* 87835 90189: contig of 2355 bp in length  
 \* 90190 90289: gap of unknown length  
 \* 90290 91900: contig of 1611 bp in length

Query Match 90.5%; Score 19; DB 2; Length 170363;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ATCATGCTTTTCTCCGG 20  
 Db 157243 ATCATGCTTTTCTCCGG 157261

RESULT 14  
 AL441943  
 LOCUS AL441943 203234 bp DNA linear PRI 15-JUL-2001  
 DEFINITION Human DNA sequence from clone RP11-69C17 on chromosome 10, complete  
 sequence.

ACCESSION AL441943  
 VERSION AL441943.6 GI:14801078  
 KEYWORDS HFG.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 203234)

REFERENCE  
 AUTHORS Lawlor, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 requests: clonerequest@sanger.ac.uk

COMMENT On Jul 17, 2001 this sequence version replaced gi:14575281.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em: EMBL; Sw:  
 SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP  
 database can be found at

http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 10, constructed by the Sanger Centre Chromosome 10  
 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr10  
 RP11-69C17 is from the library RPI1-11.1 constructed by the group  
 of Pieter de Jong. For further details see  
 http://www.chori.org/bacpac/home.htm

VECTOR: pBAC3.6

This sequence is the entire insert of clone RP11-69C17 The true  
 left end of clone RP11-362D13 is at 91113 in this sequence. The  
 true right end of clone RP11-358N3 is at 80918 in this sequence.  
 Location/Qualifiers

#### FEATURES

1. 203234  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="10"  
 /clone="RP11-69C17"  
 /clone\_lib="RPI1-11.1"  
 182..402  
 /note="MLTIE repeat: matches 288..519 of consensus"  
 471..677  
 /note="MIR repeat: matches 23..230 of consensus"  
 766..882  
 /note="3 copies 39 mer 84% conserved"  
 1929..2208  
 /note="HAL1 repeat: matches 1319..1588 of consensus"  
 2755..2884  
 /note="HAL1 repeat: matches 648..779 of consensus"  
 3389..3497  
 /note="HAL1 repeat: matches 63..166 of consensus"  
 3930..4115  
 /note="MIR repeat: matches 47..243 of consensus"  
 4403..4458  
 /note="2 copies 28 mer 98% conserved"  
 4653..4738  
 /note="MADE1 repeat: matches 1..80 of consensus"  
 4815..5363  
 /note="L1PA7 repeat: matches 5572..6143 of consensus"  
 5454..6245  
 /note="9 copies 88 mer 91% conserved"  
 7384..8045  
 /note="L1MB6 repeat: matches 5529..6179 of consensus"  
 8047..8101  
 /note="Alu repeat: matches 244..300 of consensus"  
 8212..8521  
 /note="Alu repeat: matches 2..311 of consensus"  
 8851..8946  
 /note="2 copies 48 mer 99% conserved"  
 9121..9142  
 /note="11 copies 2 mer tt 100% conserved"  
 9314..9523  
 /note="MIR repeat: matches 16..242 of consensus"  
 10623..10723  
 /note="MERSA repeat: matches 38..141 of consensus"  
 12711..12832  
 /note="61 copies 2 mer ga 59% conserved"  
 12795..12834  
 /note="10 copies 4 mer gata 92% conserved"

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region





The sequencing project is supported by Japan Science Technology Corporation (JST) and The Institute of Physical and Chemical Research (RIKEN).

FEATURES  
source

Location/Qualifiers

1..74600  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.1"  
/clone="T1866"

BASE COUNT 21480 a 16242 c 16372 g 20506 t  
ORIGIN

Query Match 87.6%; Score 18.4; DB 9; Length 74600;  
Best local Similarity 95.0%; Pred. No. 2.2e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCTGG 20  
||| ||||| ||||| ||||| |||||  
Db 22982 CAGCATGCTTTTCTCTCTGG 22963

Search completed: July 8, 2003, 03:35:03  
Job time : 235.102 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 00:43:28 ; Search time 131.941 Seconds  
(without alignments)  
358.431 Million cell updates/sec

Title: US-09-647-780A-17

Perfect score: 21

Sequence: 1 catcatgcttttctctctggg 21

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_101002.\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	21	20	AAZ28823 Rat membrane metal
C 2	19.4	92.4	2286	24	AAZ28130 Soluble secreted e
C 3	19.4	92.4	2765	20	AAZ28810 Rat membrane metal
C 4	19.4	92.4	2925	21	AAAG3763 cDNA encoding neut
C 5	17.8	84.8	1804	23	AAAG3888 DNA encoding novel
C 6	17.4	82.9	51	22	AAQ79862 Human DNA containi
C 7	17.4	82.9	329	13	AAQ28175 Human T lymphocyte
C 8	17.4	82.9	1416	24	ABL91759 Human polynucleoti
C 9	17.4	82.9	1416	24	AA44663 Human matrix metal

10	17.4	82.9	1521	20	AAZ90499 Human matrix metal
11	17.4	82.9	2698	22	AAH28230 Nucleotide sequenc
12	17.4	82.9	2753	22	AAH28230 Human cervical can
13	17.4	82.9	2792	20	AAZ90501 Tet-07/MMP13* tran
14	17.4	82.9	11495	24	ABL45714 Human matrix metal
15	16.8	80.0	406	22	AAI93579 Human polynucleoti
16	16.8	80.0	418	22	AAH98808 Human EST-derived
17	16.8	80.0	750	21	AAAS2068 Hydrophobic domain
18	16.8	80.0	1034	20	AAZ41987 Human endometrium
19	16.8	80.0	1100	19	AAV43613 Human secreted pro
20	16.8	80.0	1137	24	ABL90373 Human polynucleoti
21	16.8	80.0	1425	20	AAI19486 Human secreted pro
22	16.8	80.0	1480	20	AAV84607 Human secreted pro
23	16.8	80.0	1480	22	AAH83390 Human secreted pro
24	16.8	80.0	1493	21	AAAG2078 Hydrophobic domain
25	16.8	80.0	84607	20	AAZ90847 Human PACAP genomi
26	16.4	78.1	1632	24	ABK63615 Rat sequence diffe
27	16.4	78.1	1971	20	AAZ07301 Bacillus subtilis
28	16.4	78.1	11279	21	AAZ38389 Pseudomonas sp. WF
29	16.4	78.1	22777	22	AAK77120 Human immune/haema
30	16.4	78.1	38997	24	AAD36069 Human snare YKT6 g
31	16.2	77.1	421	24	ABN78740 Human ORF3687 cDNA
32	16.2	77.1	1175	23	AAAS85060 DNA encoding novel
33	16.2	77.1	1779	23	AAZ77191 DNA encoding novel
34	16.2	77.1	2049	23	AAZ74777 DNA encoding novel
35	16.2	77.1	3111	23	ABL13753 Drosophila melanog
36	16.2	77.1	3215	23	AAZ91206 DNA encoding novel
37	16.2	77.1	5290	22	AAK81072 Human immune/haema
38	16.2	77.1	8028	23	ABL13752 Drosophila melanog
39	16.2	77.1	8625	21	AAZ75834 Human ORF ORF1389
40	16.2	77.1	13006	22	AAK68281 Human immune/haema
41	16.2	77.1	26379	21	AAZ88922 Human wolframin ge
42	16.2	77.1	30610	22	ABAI5643 Human nervous syst
43	16.2	77.1	67212	21	AAA08954 WFS1 variant genom
44	16.2	77.1	123219	23	AAH88703 Human DNA sequence
45	16.2	77.1	305107	22	AAH62689 Shrimp white spot

#### ALIGNMENTS

RESULT 1

AAZ28823

ID AAZ28823 standard; DNA; 21 BP.

XX AAZ28823;

XX AC

DT 01-FEB-2000 (first entry)

XX DE Rat membrane metalloprotease NEPII gene probe #13.  
XX KW Rat; membrane metalloprotease; neprilysine II; NEPII; inactivation; ss;  
KW neuron; hormone; peptide messenger; inhibitor; detection; disorder;  
KW cardiovascular disease; neurodegenerative disease; growth disorder;  
KW hypothalamic-hypophyseal axis; endocrine disorder; probe; hybridisation.

XX OS Synthetic.

OS Rattus rattus.

XX PN FR2777291-Al.

XX PD 15-OCT-1999.

XX PF 08-APR-1998; 98PR-0004389.

XX PR 08-APR-1998; 98PR-0004389.

XX PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX PI Quimet T, Gros C, Haret C, Bonhomme MC, Facchinetti P;

XX PI Schwartz JC;

XX DR WPI; 1999-593429/51.

XX New membrane metalloprotease NEP II, involved in proteolysis of  
PT neuronal and hormonal peptides, used to screen for inhibitors,  
PT potentially useful for treating e.g. cardiovascular disease  
XX  
XX Claim 3; Page 23; 29pp; French.  
XX  
XX Sequences AA228811-228827 represent probes for detecting the rat  
CC membrane metalloprotease designated neprilysine II (NEPII) gene.  
CC (AA228810). NEPII is involved in (in)activation of neuronal and hormonal  
CC peptide messengers. NEPII is used to screen for specific substrates (used  
CC to detect NEPII in cells and tissues) or inhibitors, which can also be  
CC used to detect NEPII or for treatment of disorders related to peptidergic  
CC signalling in which NEPII is involved, e.g. cardiovascular or  
CC neurodegenerative diseases; growth disorders of endocrine origin;  
CC disturbances of the hypothalamic- hypophyseal axis or endocrine  
CC disorders.  
XX  
XX Sequence 21 BP; 2 A; 6 C; 4 G; 9 T; 0 other;  
SQ  
Query Match 100.0%; Score 21; DB 20; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CATCATGCTTTTCTCTCTGGG 21  
DB 1 CATCATGCTTTTCTCTCTGGG 21  
RESULT 2  
AAD28130/c  
ID AAD28130 standard; DNA; 2286 BP.  
XX  
AC AAD28130;  
XX  
DT 07-MAY-2002 (first entry)  
DE Soluble secreted endopeptidase (SEP) consensus DNA.  
XX  
KW Human; secreted endopeptidase; SEP; endocrine; vasotropic; analgesic;  
KW enzyme; gynaecological; antisense-therapy; male erectile dysfunction;  
KW MED; female sexual dysfunction; FSD; female sexual arousal disorder;  
KW FSD; premature ejaculation; anorgasmia; vaginismus; mouse; rat; ds.  
XX  
OS Homo sapiens.  
OS Mus sp.  
OS Rattus sp.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 1664..2286  
FT /\*tag= a  
FT /note= "Encodes catalytic domain"  
XX  
PN WO200206492-A1.  
XX  
PD 24-JAN-2002.  
XX  
PF 16-JUL-2001; 2001WO-1B01263.  
XX  
PR 14-JUL-2000; 2000GB-0017387.  
XX  
PA (PFIZ ) PFIZER LTD.  
PA (PFIZ ) PFIZER INC.  
XX  
PI Harrow ID, Stacey P., Walsh RT, Wayman CP, Wayman CP, Phillips SC;  
XX  
DR WPI; 2002-155042/20.  
XX  
PT An isolated and/or purified nucleic acid encoding a human soluble  
PT secreted endopeptidase which is useful for treating sexual dysfunction,  
PT for e.g. male erectile dysfunction or female sexual dysfunction such as  
PT female sexual arousal disorder  
XX

PS Disclosure; Fig 6; 167pp; English.  
XX  
CC The invention relates to an isolated and/or purified nucleic acid  
CC encoding a human soluble secreted endopeptidase (SEP). The anti-SEP  
CC antibody and the compound which inhibits or selectively inhibits the  
CC human SEP protein are useful in the manufacture of a medicament for  
CC the prophylaxis and/or treatment of sexual dysfunction, in particular  
CC male erectile dysfunction (MED) or female sexual dysfunction (FSD)  
CC (e.g. female sexual arousal disorder (FSAD)). The SEP polypeptides are  
CC also useful for treating the above disorders and other disorders such  
CC as premature ejaculation, anorgasmia, sexual pain (e.g. vaginismus) and  
CC hypoaesthetic sexual desire disorder. The present sequence is SEP consensus  
CC DNA sequence found in human, mouse and rat.  
XX  
SQ Sequence 2286 BP; 565 A; 601 C; 674 G; 446 T; 0 other;  
Query Match 92.4%; Score 19.4; DB 24; Length 2286;  
Best Local Similarity 95.2%; Pred. No. 25;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CATCATGCTTTTCTCTCTGGG 21  
DB 945 CATCATGCTTTTCTCTCTGGG 925  
RESULT 3  
AAZ28810/c  
ID AAZ28810 standard; cDNA; 2765 BP.  
XX  
AC AAZ28810;  
XX  
DT 01-FEB-2000 (first entry)  
DE Rat membrane Metalloprotease NEPII gene.  
XX  
KW Rat; membrane metalloprotease; neprilysine II; NEPII; inactivation;  
KW neuron; hormone; peptide messenger; inhibitor; detection; disorder;  
KW cardiovascular disease; neurodegenerative disease; growth disorder;  
KW hypothalamic-hypophyseal axis; endocrine disorder; ds.  
XX  
OS Rattus rattus.  
XX  
PN FR2777291-A1.  
XX  
PD 15-OCT-1999.  
XX  
PF 08-APR-1998; 98FR-0004389.  
XX  
PR 08-APR-1998; 98FR-0004389.  
XX  
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
XX  
PI Ouimet T., Gros C., Haret C., Bonhomme MC, Facchinetti P;  
PI Schwartz JC;  
XX  
XX WPI; 1999-593429/51.  
DR P-PSDB; AAY44177.  
XX  
XX New membrane metalloprotease NEP II, involved in proteolysis of  
PT neuronal and hormonal peptides, used to screen for inhibitors,  
PT potentially useful for treating e.g. cardiovascular disease  
XX  
PS Claim 2; Page 12-16; 29pp; French.  
XX  
XX This sequence represents the gene for the rat membrane metalloprotease  
CC designated neprilysine II (NEPII), which is involved in (in)activation  
CC of neuronal and hormonal peptide messengers. NEPII is used to screen  
CC for specific substrates (used to detect NEPII in cells and tissues) or  
CC inhibitors, which can also be used to detect NEPII or for treatment of  
CC disorders related to peptidergic signalling in which NEPII is involved,  
CC e.g. cardiovascular or neurodegenerative diseases; growth disorders of  
CC endocrine origin; disturbances of the hypothalamic-hypophyseal axis or  
CC endocrine disorders.  
XX

XX SQ Sequence 2765 BP; 684 A; 735 C; 787 G; 559 T; 0 other;  
Query Match 92.4%; Score 19.4; DB 20; Length 2765;  
Best Local Similarity 95.2%; Pred. No. 25;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCTGGG 21  
||||||| |||||||  
DB 1073 CATCATGCTTTTCTCTCTGGG 1053

RESULT 4  
AAA63763/C  
ID AAA63763 standard; cDNA; 2925 BP.  
XX AC AAA63763;  
XX DT 04-DEC-2000 (first entry)  
XX DE cDNA encoding neutral endopeptidase metalloproteinase-like enzyme NL-1.  
XX KW Neprilysin; neutral endopeptidase metalloproteinase-like enzyme;  
XX KW NEP-like enzyme; protein production; protein secretion;  
XX KW neurological disease; Alzheimer's disease; pain; psychiatric disorder;  
XX KW fertility; bone disease; abnormal phosphate metabolism; ss.  
XX OS Mus sp.  
XX FH Location/Qualifiers  
XX CDS 332..2629  
XX FT /\*tag= a  
XX FT /product= "neutral endopeptidase metalloproteinase-like  
XX FT enzyme NL-1"  
XX PN WO200047750-A2.  
XX PD 17-AUG-2000.  
XX PF 11-FEB-2000; 2000WO-CA00147.  
XX PR 11-FEB-1999; 99CA-2260376.  
XX PA (UYMO-) UNIV MONTREAL.  
XX PI Desgroseillers L, Boileau G;  
XX P1 WPI; 2000-549148/50.  
XX DR P-PSDB; AAB08130.  
XX CC Novel neutral endopeptidase-like metalloproteinase polypeptides and  
XX CC polynucleotides, used to screen for related sequences and enzyme  
XX CC inhibitors, used for the treatment of NL-3 related bone disorders -  
XX PS Disclosure; Fig 3; 59pp; English.  
XX CC The present sequence encodes a murine neutral endopeptidase  
XX CC metalloproteinase-like enzyme, designated NL-1. The specification  
XX CC also describes NL-2 and NL-3. The NL enzymes are used to test for  
XX CC specific inhibitors. The N-terminal region of the enzymes can be used  
XX CC to promote production and secretion of foreign proteins and active  
XX CC biopeptides, using chimeric constructs containing the foreign protein  
XX CC downstream from and in phase with the N-terminal region. The NL enzymes  
XX CC are have been localised to the brain, and may be useful in the  
XX CC treatment of neurological diseases such as Alzheimer's disease, pain,  
XX CC and psychiatric disorders. NL enzymes have also been localised to the  
XX CC testis and ovaries, and may be used to control fertility. They have  
XX CC also been localised to bones, and may be used to treat bone diseases,  
XX CC and abnormal phosphate metabolisms related to improper peptide  
XX CC processing by the NL-3 enzyme.  
XX SQ Sequence 2925 BP; 710 A; 797 C; 836 G; 582 T; 0 other;

Query Match 92.4%; Score 19.4; DB 21; Length 2925;  
Best Local Similarity 95.2%; Pred. No. 25;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCTGGG 21  
||||||| |||||||  
DB 1271 CATCATGCTTTTCTCTCTGGG 1251

RESULT 5  
AAS64388/C  
ID AAS64388 standard; cDNA; 1804 BP.  
XX AC AAS64388;  
XX DT 13-FEB-2002 (first entry)  
XX DE DNA encoding novel human diagnostic protein #192.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX OS Homo sapiens.  
XX PN WO200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US08631.  
XX PR 31-MAR-2000; 2000US-0540217.  
XX PR 23-AUG-2000; 2000US-0649167.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Drmanac RT, Liu C, Tang YT;  
XX P1 WPI; 2001-639362/73.  
XX DR P-PSDB; ABG00201.  
XX CC New isolated polynucleotide and encoded polypeptides, useful in  
XX CC diagnostics, forensics, gene mapping, identification of mutations  
XX CC responsible for genetic disorders or other traits and to assess  
XX CC biodiversity  
XX PS Claim 1; SEQ ID No 192; 103pp; English.  
XX CC The invention relates to isolated polynucleotide (I) and  
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX CC and gene mapping, and in recombinant production of (II). The  
XX CC polynucleotides are also used in diagnostics as expressed sequence tags  
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques  
XX CC to restore normal activity of (II) or to treat disease states involving  
XX CC (II). (II) is useful for generating antibodies against it, detecting or  
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as  
XX CC a food supplement. (II) and its binding partners are useful in medical  
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating  
XX CC disorders involving aberrant protein expression or biological activity.  
XX CC The polypeptide and polynucleotide sequences have applications in  
XX CC diagnostics, forensics, gene mapping, identification of mutations  
XX CC responsible for genetic disorders or other traits to assess biodiversity  
XX CC and to produce other types of data and products dependent on DNA and  
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human  
XX CC diagnostic coding sequences of the invention.  
XX CC Note: The sequence data for this patent did not appear in the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 1804 BP; 441 A; 465 C; 433 G; 465 T; 0 other;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCTGGG 21  
 |||| ||||| ||||| |||||  
 Db 1255 CATCTGCTTTGTCTCTCTGGG 1235

RESULT 6  
 AAH79862  
 ID AAH79862 standard; DNA; 51 BP.  
 XX  
 AC AAH79862;  
 XX  
 DT 19-SEP-2001 (first entry)  
 XX  
 DE Human DNA containing single nucleotide polymorphism SEQ ID NO. 477.  
 XX  
 KW Human; single nucleotide polymorphism; SNP; angiotensinogen;  
 KW 4-hydroxybutyrate; dehydrogenase; protein therapy;  
 KW adenosine triphosphate-dependent RNA helicase;  
 KW major histocompatibility complex Class I histocompatibility antigen; MHC;  
 KW phosphoglycerate kinase; immunosuppressive; immunostimulatory;  
 KW antirheumatic; antisclerotic; antidiabetic; antiinflammatory; cytostatic;  
 KW antileukemic; neuroprotective; antimicrobial; gene therapy; vaccine; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200148245-A2.  
 XX  
 PD 05-JUL-2001.  
 XX  
 PF 27-DEC-2000; 2000WO-0535346.  
 XX  
 PR 27-DEC-1999; 99US-0472688.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shinkets RA, Leach M;  
 XX  
 DR WPI; 2001-418297/44.  
 XX  
 XX Polymorphic nucleic acids encoding e.g. angiotensinogen, dehydrogenase,  
 PT adenosine triphosphate-dependent RNA helicase and/or phosphoglycerate  
 PT kinase, useful for diagnosing and treating, e.g. cancer, autoimmune  
 PT diseases and infections  
 XX  
 PS Claim 1; Page 193; 484pp; English.  
 XX  
 CC The invention relates to nucleic acids (AAH79386-AAH80036) encoding  
 CC polymorphic variants of proteins (AAG98010-AAG98238) related to  
 CC angiotensinogen, 4-hydroxybutyrate, dehydrogenase, adenosine triphosphate  
 CC (ATP)-dependent RNA helicase, major histocompatibility complex (MHC)  
 CC Class I histocompatibility antigen and/or phosphoglycerate kinase. These  
 CC nucleic acid single nucleotide polymorphisms (SNPs) and the encoded  
 CC proteins have potential immunosuppressive, immunostimulatory,  
 CC antirheumatic, antisclerotic, antidiabetic, antiinflammatory, cytostatic,  
 CC antileukemic, neuroprotective and antimicrobial activity and may be  
 CC useful in gene/protein therapy, vaccines, modulation of the expression  
 CC and activity of proteins related to angiotensinogen, 4-hydroxybutyrate,  
 CC dehydrogenase, adenosine triphosphate (ATP)-dependent RNA helicase,  
 CC major histocompatibility complex (MHC) Class I histocompatibility antigen  
 CC diagnosed and/or treated by the above methods that may be prevented,  
 CC diseases with a genetic component, such as autoimmune diseases (e.g.  
 CC rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus  
 CC erythematosus and Grave's disease), inflammation, cancer (e.g. cancers  
 CC of the bladder, brain, breast, colon and kidney, leukemia), diseases of  
 CC the nervous system, an infection of pathogenic organisms. They may also  
 CC be used to alter phenotypic traits such as longevity, appearance,  
 CC strength, speed and endurance.  
 XX  
 SQ Sequence 51 BP; 8 A; 17 C; 13 G; 13 T; 0 other;

Query Match 82.9%; Score 17.4; DB 22; Length 51;  
 Best Local Similarity 94.7%; Pred. No. 1.2e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCATGCTTTTCTCTCTGGG 21  
 ||||| ||||| ||||| |||||  
 Db 29 TCATGCTTTTCTCTCTGGG 47

RESULT 7  
 AAQ28175  
 ID AAQ28175 standard; cDNA to mRNA; 329 BP.  
 XX  
 AC AAQ28175;  
 XX  
 DT 11-FEB-1993 (first entry)  
 XX  
 DE Human T lymphocyte receptor V-region "IGR b 04".  
 XX  
 KW TCR; V beta w23 subfamily; variable region; anchored PCR;  
 KW polymerase chain reaction; T cell receptor; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT mat\_peptide 18..329  
 FT /\*tag= a  
 FT /note= "encodes a V beta segment"  
 XX  
 PN WO9213950-A.  
 XX  
 PD 20-AUG-1992.  
 XX  
 PF 12-FEB-1992; 92WO-FR00130.  
 XX  
 PR 12-FEB-1991; 91FR-0001613.  
 XX  
 PR 12-APR-1991; 91FR-0004523.  
 XX  
 PA (ROUS) ROUSSEL-UCILAF.  
 XX  
 PI Ferradini L, Hercend T, Roman-Roman S, Triebel F;  
 XX  
 DR WPI; 1992-300036/36.  
 XX  
 DR P-PSDB; AAR26962.  
 XX  
 PT Variable regions of b-chain of T-lymphocyte receptors and their  
 PT DNA - useful as immuno:modulant(s) and for diagnosing immune  
 PT disorders  
 XX  
 PS Claim 1; Page 39; 75pp; French.  
 XX  
 CC RNA was isolated from peripheral lymphocytes and converted to cDNA  
 CC using a C-beta-specific primer. The cDNA was amplified by anchored  
 CC PCR using C-beta and poly C primers, then amplified again using a  
 CC different C-beta specific primer. The amplified product was SacII-  
 CC restricted, inserted into Bluescript SK+ vector and used to transform  
 CC E.coli XL-blue. transformants were screened with a C-beta specific  
 CC probe and DNA from positive clones was sequenced in the C-beta  
 CC region. The sequence designated "IGR b 04" is a consensus sequence  
 CC from 4 distinct cDNA clones; a G was seen at position 154 in place  
 CC of a A and an A at position 160 instead of a G. The sequence has an  
 CC homology of 75.7% with the sequence Vbeta1A1 (see Leiden J.M., et  
 CC al., Proc.Natl.Acad.Sci. USA, 83:4456, 1986) but has a homology of  
 CC less than 75% with other members of the Vbeta 5 subfamily. IGR b 04  
 CC is, therefore, not a member of the Vbeta 5 subfamily.  
 CC See AAQ28173-Q28228.  
 XX  
 SQ Sequence 329 BP; 83 A; 92 C; 75 G; 79 T; 0 other;

Query Match 82.9%; Score 17.4; DB 13; Length 329;  
 Best Local Similarity 94.7%; Pred. No. 1.6e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



OY 3 TCATGCTTTTCTCTCTGGG 21  
Db 16 TCATGCTTTTCTCTCTGGG 34

RESULT 8  
ABL91759  
ID ABL91759 standard; DNA; 1416 BP.

XX  
AC ABL91759;  
XX  
DT 28-MAY-2002 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 102.  
XX  
KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;  
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;  
KW Cytostatic; virucide; protozoacide; antibacterial; ds.  
XX  
OS Homo sapiens.  
XX  
PN DE10100586-C1.  
XX  
PD 11-APR-2002.  
XX  
PF 09-JAN-2001; 2001DE-1000586.  
XX  
PR 09-JAN-2001; 2001DE-1000586.  
XX  
PA (RIBO-) RIBOPHARMA AG.  
XX  
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;  
XX  
DR WPI; 2002-270454/32.  
XX  
PT Inhibiting gene expression in cells, useful for e.g. treating tumors,  
PT by introducing double-stranded complementary oligoRNA having unpaired  
PT terminal bases  
XX  
PS Claim 13; Page 75-76; 104pp; German.

The invention relates to a method for inhibiting expression of a target gene (ABL91658-ABL91797) in a cell by introducing at least one oligoribonucleotide that has a double-stranded structure consisting of at most 49 sequential nucleotide pairs, with at least part of one strand complementary with the target gene and has at least one end a single-stranded segment of 1-4 nt. The method provides oligoribonucleotides for antisense inhibition of gene expression useful e.g. for treating tumors but the oligoribonucleotides may also be directed against genes present in pathogens (e.g. Plasmodium or viruses/viroids, pathogenic on humans, animals or plants) or against cytokine, id, developmental or prion-genes. The method provides more effective inhibition of gene expression than use of known oligonucleotides, probably because the unpaired overhang increases stability and thus intracellular concentration.

XX  
SQ Sequence 1416 BP; 389 A; 328 C; 306 G; 393 T; 0 other;  
Query Match 82.9%; Score 17.4; DB 24; Length 1416;  
Best Local Similarity 94.7%; Pred. No. 1.9e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 TCATGCTTTTCTCTCTGGG 21  
Db 558 TCATGCTTTTCTCTCTGGG 576

RESULT 9  
AAL44663  
ID AAL44663 standard; cDNA; 1416 BP.  
XX  
AC AAL44663;  
XX

DT 03-MAY-2002 (first entry)  
XX  
DE Human matrix metalloproteinase 13 (collagenase 3) cDNA #2.  
XX  
KW Human; matrix metalloproteinase 13 (collagenase 3); MMP13; cancer;  
KW arthritis; haplotype; single nucleotide polymorphism; SNP; enzyme;  
KW Cytostatic; antiarthritic; gene therapy; chromosome 11q22.3; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT 1..1416  
FT /\*tag= a  
FT /product= "MMP13"  
FT replace(326,T)  
FT allele  
FT /\*tag= b  
FT replace(1080,C)  
FT allele  
FT /\*tag= c  
FT replace(1169,G)  
FT allele  
FT /\*tag= d  
XX  
PN WO200206294-A2.  
XX  
PD 24-JAN-2002.  
XX  
PF 13-JUL-2001; 2001WO-US22238.  
XX  
PR 13-JUL-2000; 2000US-217950P.  
PR 17-AUG-2000; 2000WO-US22693.  
XX  
PA (GENA-) GENAISANCE PHARM INC.  
XX  
PI Finkel K, Kliem SE, Messer C, Tanguay DA;  
XX  
DR WPI; 2002-171797/22.  
XX  
PT Novel genetic variants of matrix metalloproteinase 13 (collagenase 3)  
PT gene useful in studying expression and function of the protein, and for  
PT screening drugs to treat diseases e.g. cancer and arthritis  
XX  
PS Claim 25; Fig 4; 110pp; English.

The present invention provides the cDNA, protein and gene fragments of the human matrix metalloproteinase 13 (collagenase 3) (MMP13). Also provided are single nucleotide polymorphisms (SNPs) identified within the sequences. The sequences can be used to haplotype an individual and in the treatment of cancer and arthritis, including metastatic cancers. The present sequence is one version of the MMP13 cDNA, the gene for which is found on chromosome 11q22.3.  
CC Note: The present sequence is stated as being the same as that shown as  
CC SEQ ID NO: 4 in the sequence listing of the specification. However, the  
CC sequences differ.  
XX  
SQ Sequence 1416 BP; 388 A; 328 C; 307 G; 393 T; 0 other;  
Query Match 82.9%; Score 17.4; DB 24; Length 1416;  
Best Local Similarity 94.7%; Pred. No. 1.9e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 TCATGCTTTTCTCTCTGGG 21  
Db 558 TCATGCTTTTCTCTCTGGG 576

RESULT 10  
AAX90499  
ID AAX90499 standard; DNA; 1521 BP.  
XX  
AC AAX90499;  
XX  
DT 04-OCT-1999 (first entry)  
XX

DE XX Human matrix metalloproteinase 13 encoding DNA.  
KW Human; matrix metalloproteinase 13; MMP-13; collagenase 3; cartilage;  
KW degenerative disease; matrix degrading enzyme; MDE; transgenic mammal;  
KW ss.  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 5..1420  
FT /\*tag= a  
FT /product= "MMP-13"  
FT /note= "matrix metalloproteinase"  
XX  
XX WO9931969-A2.  
XX  
XX 01-JUL-1999.  
XX  
XX 18-DEC-1998; 98WO-US27056.  
XX 19-DEC-1997; 97US-0994689.  
XX 19-DEC-1997; 97US-0068312.  
XX (AMHP ) AMERICAN HOME PROD. CORP.  
XX  
XX Killar LM, Neuhold LA;  
PI  
XX  
XX WPI; 1999-468690/39.  
XX P-PSDB; AAY29419.  
XX  
XX Transgenic mammals that express recombinant matrix-degrading  
PT enzymes, used to study phenotypic changes associated with  
PT cartilage-degenerative disease  
XX  
XX Claim 7; Fig 1B-1C; 70pp; English.  
XX  
XX The present invention describes transgenic mammals that express  
CC recombinant matrix-degrading enzymes (MDE) in a temporally and spatially  
CC regulated manner. The transgenic mammals produce phenotypic changes  
CC associated with cartilage-degenerative disease if maintained under  
CC conditions in which the recombinant gene is selectively expressed in  
CC joint tissue. The transgenic animal models are useful for determining  
CC the potential of a composition to counteract cartilage-degenerative  
CC disease. The present sequence encodes human matrix metalloproteinase 13  
CC (MMP-13) which can be used as the recombinant MDE in a transgenic  
CC mammal of the present invention.  
XX  
XX Sequence 1521 BP; 422 A; 339 C; 338 G; 422 T; 0 other;  
SQ  
Query Match 82.9%; Score 17.4; DB 20; Length 1521;  
Best Local Similarity 94.7%; Pred. No. 1.9e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3 TCATGCTTTTCTCCTGGG 21  
|||||||  
Db 562 TCATGCTTTTCTCCTGGG 580  
RESULT 11  
AAH28230  
ID AAH28230 standard; cDNA; 2698 BP.  
XX  
XX AAH28230;  
XX  
XX 05-SEP-2001 (first entry)  
XX  
XX Nucleotide sequence of matrix metalloproteinase-13.  
DE  
XX Growth factor; protein inhibitor; protease; damaged tissue;  
KW platelet-derived growth factor; PDGF; fibroblast growth factor; FGF;  
KW connective tissue derived growth factor; CTGF; chrysalin; VEGF;  
KW keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;  
KW transforming growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;

KW granulocyte macrophage colony stimulating factor; GM-CSF; uPA;  
KW vascular endothelial growth factor; urokinase plasminogen activator;  
KW dermal ulcer; wound; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 5..1420  
FT /\*tag= a  
FT /product= "MMP-13"  
FT /note= "matrix metalloproteinase"  
XX  
XX WO200149309-A2.  
XX  
XX 12-JUL-2001.  
XX  
XX 21-DEC-2000; 2000WO-IB01935.  
XX  
XX 29-DEC-1999; 99GB-0030768.  
XX (PFIZ ) PFIZER LTD.  
XX (PFIZ ) PFIZER INC.  
XX  
XX Davies MJ, Huggins JP, McIntosh FS, Occleston NL;  
PI  
XX WPI; 2001-418351/44.  
XX P-PSDB; AAB84616.  
XX  
XX Composition for the treatment of damaged tissue i.e. chronic wounds and  
PT dermal ulcers comprises an inhibitor agent i.e. a protease and a growth  
PT factor  
XX  
XX Disclosure; Page 561-562; 572pp; English.  
XX  
XX The specification describes a pharmaceutical composition, comprising  
CC a growth factor, an inhibitor agent, i.e. a protease. The inhibitor  
CC agent inhibits the action of at least one specific adverse protein,  
CC i.e. a protease, that is upregulated in a damaged tissue such as a  
CC wound environment. Growth factors which are included in the composition  
CC of the invention are platelet-derived growth factor (PDGF), fibroblast  
CC growth factor (FGF), connective tissue derived growth factor (CTGF),  
CC keratinocyte-derived growth factor (KGF), transforming growth  
CC factor-beta (TGF-beta), granulocyte macrophage colony stimulating factor  
CC (GM-CSF), epidermal growth factor (EGF), vascular endothelial growth  
CC factor (VEGF), and chrysalin. Inhibitors which are included in the  
CC composition of the invention include inhibitors of urokinase-type  
CC plasminogen activator (uPA) and matrix metalloproteinase (MMP). The  
CC composition is useful for the treatment of chronic damaged tissue, i.e.  
CC wounds and dermal ulcers. The present sequence encodes a human  
CC MMP-13, and is used to produce the composition of the invention.  
XX  
XX Sequence 2698 BP; 853 A; 505 C; 540 G; 800 T; 0 other;  
SQ  
Query Match 82.9%; Score 17.4; DB 22; Length 2698;  
Best Local Similarity 94.7%; Pred. No. 2.1e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3 TCATGCTTTTCTCCTGGG 21  
|||||||  
Db 562 TCATGCTTTTCTCCTGGG 580  
RESULT 12  
AAH72608  
ID AAH72608 standard; cDNA; 2753 BP.  
XX  
XX AAH72608;  
XX  
XX 19-SEP-2001 (first entry)  
XX  
XX Human cervical cancer marker nucleic acid 3882.  
DE  
XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.  
KW  
XX

```
OS Homo sapiens.
XX WO200142467-A2.
XX
XX
XX
XX 14-JUN-2001.
XX
XX 08-DEC-2000; 2000WO-US33312.
XX
XX 08-DEC-1999; 99US-0169681.
XX 21-DEC-1999; 99US-0171350.
XX 14-MAR-2000; 2000US-0169315.
XX 12-MAY-2000; 2000US-0203791.
XX 09-JUN-2000; 2000US-0210600.
XX 21-JUL-2000; 2000US-0220114.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Deeds J, Berger A, Zhao X;
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer
XX and for assessing and detecting compounds for treating the cancer
XX
XX Claim 1; Page 757-758; 1051pp; English.
XX
XX The invention relates to novel genes (AAH68727-AAH73383) associated with
XX cervical cancer with cytostatic activity. The nucleic acids and encoded
XX polypeptides are useful: to assess if a patient is afflicted with
XX cervical cancer or has a pre-malignant condition; to monitor the
XX progression of cervical cancer or a premalignant condition in a patient;
XX and to select and/or assess the efficacy of a compound or therapy for
XX inhibiting cervical cancer in a patient. The nucleic acids may also be
XX useful for gene therapy.
XX
XX Sequence 2753 BP; 857 A; 523 C; 557 G; 806 T; 10 other;
XX
Query Match 82.9%; Score 17.4; DB 22; Length 2753;
Best Local Similarity 94.7%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
OY 3 TCATGCTTTTCTCCTGGG 21
DB 615 TCATGCTTTTCTCCTGGG 633
XXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXX
RESULT 13
AAX90501
ID AAX90501 standard; DNA; 2792 BP.
XX
XX AAX90501;
XX
XX 04-OCT-1999 (first entry)
XX
XX Tet-07/MMP13* transgene.
XX
XX Human; matrix metalloproteinase 13; MMP-13; collagenase 3; cartilage;
XX degenerative disease; matrix degrading enzyme; MDE; transgenic mammal;
XX ss.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX WO9931969-A2.
XX
XX 01-JUL-1999.
XX
XX 18-DEC-1998; 98WO-US27056.
XX
XX 19-DEC-1997; 97US-0994689.
XX 19-DEC-1997; 97US-0068312.
XX
XX (AMHP ) AMERICAN HOME PROD CORP.
XX PA
```

```
XX Killar LM, Neuhold LA;
XX WPI; 1999-468690/39.
XX
XX Transgenic mammals that express recombinant matrix-degrading
XX enzymes, used to study phenotypic changes associated with
XX cartilage-degenerative disease
XX
XX Example 3; Page 63-64; 70pp; English.
XX
XX The present invention describes transgenic mammals that express
XX recombinant matrix-degrading enzymes (MDE) in a temporally and spatially
XX regulated manner. The transgenic mammals produce phenotypic changes
XX associated with cartilage-degenerative disease if maintained under
XX conditions in which the recombinant gene is selectively expressed in
XX joint tissue. The transgenic animal models are useful for determining
XX the potential of a composition to counteract cartilage-degenerative
XX disease. The present sequence represents a transgene with a tet-07
XX promoter driving expression of a constitutively active human MMP-13
XX protein followed by an SV40 splice and polyadenylation signal, where
XX the transgene is designated Tet-07/MMP13*.
XX
XX Sequence 2792 BP; 814 A; 598 C; 601 G; 779 T; 0 other;
XX
Query Match 82.9%; Score 17.4; DB 20; Length 2792;
Best Local Similarity 94.7%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
OY 3 TCATGCTTTTCTCCTGGG 21
DB 1059 TCATGCTTTTCTCCTGGG 1077
XXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXX
RESULT 14
ABL45714
ID ABL45714 standard; DNA; 11495 BP.
XX
XX ABL45714;
XX
XX 03-MAY-2002 (first entry)
XX
XX Human matrix metalloproteinase 13 (collagenase 3) gene fragment #1.
XX
XX Human; matrix metalloproteinase 13 (collagenase 3); MMP13; cancer;
XX arthritis; haplotype; single nucleotide polymorphism; SNP; enzyme;
XX cytostatic; antiarthritic; gene therapy; chromosome 11q22.3; gene; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX allele replace(3915,G)
XX CDS /*tag= a
XX /*tag= b
XX /*product= "MMP13 fragment"
XX /*partial
XX /*note= "this sequence contains introns. no stop codon"
XX exon 4020..4139
XX /*tag= c
XX /*number= "1"
XX intron 4140..4231
XX /*tag= d
XX exon /*number= "1"
XX /*tag= e
XX /*number= "2"
XX allele replace(4437,T)
XX intron /*tag= f
XX /*number= "1"
XX /*tag= g
XX /*number= "2"
XX allele replace(5008,T)
```

FT FT allele /\*tag= h  
FT FT replace(5037,G)  
FT FT /\*tag= i  
FT FT replace(5102,G)  
FT FT /\*tag= j  
FT FT exon 5119..5266  
FT FT /\*tag= k  
FT FT /number= "3"  
FT FT intron 5267..5442  
FT FT /\*tag= l  
FT FT /number= "3"  
FT FT replace(5290,G)  
FT FT /\*tag= m  
FT FT allele replace(5363,A)  
FT FT /\*tag= n  
FT FT exon 5443..5569  
FT FT /\*tag= o  
FT FT /number= "4"  
FT FT intron 5570..7551  
FT FT /\*tag= p  
FT FT /number= "4"  
FT FT allele replace(5628,G)  
FT FT /\*tag= q  
FT FT exon 7552..7715  
FT FT /\*tag= r  
FT FT /number= "5"  
FT FT intron 7716..9499  
FT FT /\*tag= s  
FT FT /number= "5"  
FT FT allele replace(7721,G)  
FT FT /\*tag= t  
FT FT exon 9500..9615  
FT FT /\*tag= u  
FT FT /number= "6"  
FT FT intron 9616..10565  
FT FT /\*tag= v  
FT FT /number= "6"  
FT FT allele replace(10537,T)  
FT FT /\*tag= w  
FT FT exon 10566..10699  
FT FT /\*tag= x  
FT FT /number= "7"  
XX XX  
XX XX WO200206294-A2.  
XX XX  
XX XX 24-JAN-2002.  
XX XX  
XX XX 13-JUL-2001; 2001WO-US22238.  
XX XX  
XX XX 13-JUL-2000; 2000US-217950P.  
XX XX 17-AUG-2000; 2000WO-US22693.  
XX XX  
XX XX (GENA-) GENAISSANCE PHARM INC.  
XX XX  
XX XX Finkel K, Kliem SE, Messer C, Tanguay DA;  
XX XX  
XX XX WPI; 2002-171797/22.  
XX XX P-PSDB; AAM48977.  
XX XX  
XX XX Novel genetic variants of matrix metalloproteinase 13 (collagenase 3)  
XX XX gene useful in studying expression and function of the protein, and for  
XX XX screening drugs to treat diseases e.g. cancer and arthritis -  
XX XX  
XX XX Claim 20; Fig 1; 110pp; English.  
XX XX  
XX XX The present invention provides the cDNA, protein and gene fragments of  
XX XX the human matrix metalloproteinase 13 (collagenase 3) (MMP13). Also  
XX XX provided are single nucleotide polymorphisms (SNPs) identified within the  
XX XX sequences. The sequences can be used to haplotype an individual and in  
XX XX the treatment of cancer and arthritis, including metastatic cancers. The  
XX XX present sequence is a fragment of the MMP13 gene, which is found on  
XX XX chromosome 11q22.3.

SQ Sequence 11495 BP; 3577 A; 2093 C; 1948 G; 3803 T; 74 other;  
Query Match 82.9%; Score 17.4; DB 24; Length 11495;  
Best Local Similarity 94.7%; Pred. No. 2.5e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3 TCATGCTTTTCTCTCTGGG 21  
||||||| |||||||  
DB 5490 TCATGCTTTTCTCTCTGGG 5508  
RESULT 15  
AAI93579/c  
ID AAI93579 standard; cDNA; 406 BP.  
XX AC AAI93579;  
XX 06-NOV-2001 (first entry)  
XX Human polynucleotide SEQ ID NO 13639.  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
XX tissue growth factor; immunomodulatory; cancer; leukaemia;  
XX nervous system disorders; arthritis; inflammation; ss.  
XX Homo sapiens.  
XX OS  
XX WO200164835-A2.  
XX PN  
XX 07-SEP-2001.  
XX PD  
XX  
XX 26-FEB-2001; 2001WO-US04927.  
XX PF  
XX  
XX 28-FEB-2000; 2000US-0515126.  
XX PR  
XX 18-MAY-2000; 2000US-0577409.  
XX PA (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Drmanac RT;  
XX WPI; 2001-514838/56.  
XX P-PSDB; AAO13648.  
XX  
XX Isolated nucleic acids and polypeptides, useful for preventing  
XX diagnosing and treating e.g. leukaemia, inflammation and immune  
XX disorders  
XX  
XX Claim 1; SEQ ID NO 13639; 1399pp + Sequence Listing; English.  
XX  
XX The invention relates to human polynucleotides (AAI93579) and  
XX the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
XX cytokine, cell proliferation or cell differentiation or which may induce  
XX production of other cytokines in other cell populations. The  
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or  
XX peptide therapy. The polypeptides have various cytokine-like activities,  
XX e.g. stem cell growth factor activity, haematopoiesis regulating  
XX activity, tissue growth factor activity, immunomodulatory activity and  
XX activin/inhibin activity and may be useful in the diagnosis and/or  
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and  
XX inflammation.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 406 BP; 160 A; 62 C; 83 G; 101 T; 0 other;  
Query Match 80.0%; Score 16.8; DB 22; Length 406;  
Best Local Similarity 90.0%; Pred. No. 3e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CATCATGCTTTTCTCTCTGGG 20  
||||||| |||||||

Db 116 CATCATGTTTTTCCTCCGG 97

Search completed: July 8, 2003, 02:19:10  
Job time : 133.941 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 01:24:03 ; Search time 28.9902 Seconds  
(without alignments)  
222.151 Million cell updates/sec

Title: US-09-647-780A-17

Perfect score: 21  
Sequence: 1 catcatgctttctctctggg 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/lna/5A-COMB.seq:\*  
2: /cgn2\_6/ptodata/1/lna/5B-COMB.seq:\*  
3: /cgn2\_6/ptodata/1/lna/6A-COMB.seq:\*  
4: /cgn2\_6/ptodata/1/lna/6B-COMB.seq:\*  
5: /cgn2\_6/ptodata/1/lna/PTUS-COMB.seq:\*  
6: /cgn2\_6/ptodata/1/lna/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17.4	82.9	329	1	US-08-423-383-4
2	17.4	82.9	329	2	US-08-437-353A-4
3	15.8	75.2	226	4	US-09-404-879A-79
4	15.8	75.2	226	4	US-09-404-879A-253
5	15.8	75.2	226	4	US-09-404-879A-254
6	15.4	73.3	513	3	US-08-714-918-49
7	15.4	73.3	513	4	US-09-265-315-49
8	15.4	73.3	513	4	US-09-265-315-49
9	15.4	73.3	513	4	US-09-266-417-49
10	15.4	73.3	741	4	US-09-370-838-103
11	15.4	73.3	1927	4	US-08-969-046-15
12	15.4	73.3	5631	4	US-09-052-469-1
13	15.4	73.3	12912	2	US-08-460-751-1
14	15.4	73.3	13807	4	US-09-052-469-5
15	15.4	73.3	14060	3	US-08-658-136-4
16	15.4	73.3	14148	4	US-09-052-469-7
17	15.4	73.3	53526	3	US-08-658-136-2
18	15.4	73.3	53577	3	US-08-658-136-1
19	15.4	73.3	111282	4	US-09-754-250-3
20	15.4	73.3	162450	4	US-09-345-882-1
21	15.2	72.4	732	1	US-08-361-395-2
22	15.2	72.4	1072	4	US-09-280-116-212
23	15.2	72.4	2799	1	US-08-212-188-1
24	15.2	72.4	2799	3	US-08-970-725-1
25	15.2	72.4	2799	5	PCT-US95-02708-1
26	15	71.4	494	4	US-08-477-928A-35
27	15	71.4	2107	4	US-08-477-928A-3

Sequence 1, Appli  
Sequence 6, Appli  
Sequence 3, Appli  
Sequence 33, Appli  
Sequence 33, Appli  
Sequence 115, App  
Patent No. 5432264  
Sequence 32, Appli  
Sequence 104, App  
Sequence 1, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 2, Appli  
Sequence 97, Appli  
Sequence 97, Appli  
Sequence 97, Appli

Sequence 1, Appli  
Sequence 6, Appli  
Sequence 3, Appli  
Sequence 33, Appli  
Sequence 33, Appli  
Sequence 115, App  
Patent No. 5432264  
Sequence 32, Appli  
Sequence 104, App  
Sequence 1, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 2, Appli  
Sequence 97, Appli  
Sequence 97, Appli  
Sequence 97, Appli

Sequence 1, Appli  
Sequence 6, Appli  
Sequence 3, Appli  
Sequence 33, Appli  
Sequence 33, Appli  
Sequence 115, App  
Patent No. 5432264  
Sequence 32, Appli  
Sequence 104, App  
Sequence 1, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 2, Appli  
Sequence 97, Appli  
Sequence 97, Appli  
Sequence 97, Appli

## ALIGNMENTS

RESULT 1  
US-08-423-383-4  
; Sequence 4, Application US/08423383  
; Patent No. 5700907  
; GENERAL INFORMATION:  
; APPLICANT: HERCEND, THIERRY; TRIEBEL, FREDERIC;  
; APPLICANT: ROMAN-ROMAN, SERGIO; FERRADINI, LAURENT  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR  
; TITLE OF INVENTION: VARIABLE REGIONS OF BETA CHAINS OF HUMAN T LYMPHOCYTE  
; TITLE OF INVENTION: RECEPTORS, CORRESPONDING PEPTIDE SEGMENTS AND THE  
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USES  
; NUMBER OF SEQUENCES: 87  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIERMAN & MUSERLIAN  
; STREET: 600 THIRD AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/423,383  
; FILING DATE: 14-APR-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/934,530  
; FILING DATE: 23-NOV-1992  
; APPLICATION NUMBER: PCT/FR92/00130  
; FILING DATE: 12-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR/91/01613  
; FILING DATE: 12-FEB-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR/91/04523  
; FILING DATE: 12-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CHARLES A. MUSERLIAN  
; REGISTRATION NUMBER: 19,683  
; REFERENCE/DOCKET NUMBER: 146.1158  
; TELEPHONE: 212-661-8000  
; TELEFAX: 212-661-8002  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 329  
; TYPE: NUCLEOTIDE



STRANDEDNESS: DOUBLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: CDNA TO mRNA  
ORIGINAL SOURCE:  
ORGANISM: HUMAN  
CELL LINE: HUMAN T LYMPHOCYTE  
FEATURE:  
NAME/KEY: IGR b 04  
OTHER INFORMATION: V BETA W23  
US-08-437-383-4

Query Match 82.9%; Score 17.4; DB 1; Length 329;  
Best Local Similarity 94.7%; Pred. No. 22;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TCATGCTTTTCTCTCTGGG 21  
||||||| |||||  
Db 16 TCATGCTTTGTCTCTCTGGG 34

RESULT 2  
US-08-437-353A-4  
; Sequence 4, Application US/08437353A  
; Patent No. 5830758  
; GENERAL INFORMATION:  
; APPLICANT: HERCEND, THIERRY; TRIEBEL, FREDERIC;  
; APPLICANT: ROMAN-ROMAN, SERGIO; FERRADINI, LAURENT  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR  
; TITLE OF INVENTION: VARIABLE REGIONS OF BETA CHAINS OF HUMAN T LYMPHOCYTE  
; TITLE OF INVENTION: RECEPTORS, CORRESPONDING PEPTIDE SEGMENTS AND THE DIAGNOSTIC  
; TITLE OF INVENTION: AND THERAPEUTIC USES  
; NUMBER OF SEQUENCES: 87  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIERMAN & MUSERLIAN  
; STREET: 600 THIRD AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/437,353A  
; FILING DATE: 09-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/423,383  
; FILING DATE: 14-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/934,530  
; FILING DATE: 23-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR92/00130  
; FILING DATE: 12-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR/91/01613  
; FILING DATE: 12-FEB-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR/91/04523  
; FILING DATE: 12-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CHARLES A. MUSERLIAN  
; REGISTRATION NUMBER: 19,683  
; REFERENCE/DOCKET NUMBER: 146,1158  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-661-8000  
; TELEFAX: 212-661-8002  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 329

TYPE: NUCLEOTIDE  
STRANDEDNESS: DOUBLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: CDNA TO mRNA  
ORIGINAL SOURCE:  
ORGANISM: HUMAN  
CELL LINE: HUMAN T LYMPHOCYTE  
FEATURE:  
NAME/KEY: IGR b'04  
OTHER INFORMATION: V BETA W23  
US-08-437-353A-4

Query Match 82.9%; Score 17.4; DB 2; Length 329;  
Best Local Similarity 94.7%; Pred. No. 22;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TCATGCTTTTCTCTCTGGG 21  
||||||| |||||  
Db 16 TCATGCTTTGTCTCTCTGGG 34

RESULT 3  
US-09-404-879A-79  
; Sequence 79, Application US/09404879A  
; Patent No. 6468546  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: King, Gordon E.  
; APPLICANT: Algate, Paul A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.462C2  
; CURRENT APPLICATION NUMBER: US/09/404,879A  
; CURRENT FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 393  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 79  
; LENGTH: 226  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-404-879A-79

Query Match 75.2%; Score 15.8; DB 4; Length 226;  
Best Local Similarity 89.5%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CATCATGCTTTTCTCTCTG 19  
|| ||||| |||||  
Db 28 CAGCATGCTTTCTCTCTG 46

RESULT 4  
US-09-404-879A-253/C  
; Sequence 253, Application US/09404879A  
; Patent No. 6468546  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: King, Gordon E.  
; APPLICANT: Algate, Paul A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.462C2  
; CURRENT APPLICATION NUMBER: US/09/404,879A  
; CURRENT FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 393  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 253  
; LENGTH: 226  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-404-879A-253

Query Match 75.2%; Score 15.8; DB 4; Length 226;

Best Local Similarity: 89.5%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTG 19  
|| ||||| ||||| |||||  
Db 199 CAGCATGCTTTTCTCTG 181

## RESULT 5

US-09-404-879A-254  
; Sequence 254, Application US/09404879A  
; Patent No. 6468546  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: King, Gordon E.  
; APPLICANT: Algate, Paul A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.462C2  
; CURRENT APPLICATION NUMBER: US/09/404,879A  
; CURRENT FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 393  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 254  
; LENGTH: 226  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-404-879A-254

Query Match 75.2%; Score 15.8; DB 4; Length 226;  
Best Local Similarity 89.5%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTG 19  
|| ||||| ||||| |||||  
Db 28 CAGCATGCTTTTCTCTG 46

## RESULT 6

US-08-714-918-49  
; Sequence 49, Application US/08714918  
; Patent No. 6037123  
; GENERAL INFORMATION:  
; APPLICANT: Benton, Bret  
; APPLICANT: Lee, Ving  
; APPLICANT: Malouin, Francois  
; APPLICANT: Martin, Patrick K.  
; APPLICANT: Schmid, Molly B.  
; APPLICANT: Sun, Dongxu  
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL  
; TITLE OF INVENTION: TARGET GENES  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/714,918  
; FILING DATE: September 13, 1996  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/009,102  
; FILING DATE: December 22, 1995

APPLICATION NUMBER: 60/003,798  
FILING DATE: September 15, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 222/005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 513 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-714-918-49

Query Match 73.3%; Score 15.4; DB 3; Length 513;  
Best Local Similarity 88.9%; Pred. No. 1.9e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATCATGCTTTTCTCTG 19  
||||| ||||| |||||  
Db 386 ATCATGCTTTTCTCTG 403

## RESULT 7

US-09-265-315-49  
; Sequence 49, Application US/09265315  
; Patent No. 6187541  
; GENERAL INFORMATION:  
; APPLICANT: Benton, Bret  
; APPLICANT: Lee, Ving J.  
; APPLICANT: Malouin, Francois  
; APPLICANT: Martin, Patrick K.  
; APPLICANT: Schmid, Molly B.  
; APPLICANT: Sun, Dongxu  
; TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS  
; TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS  
; TITLE OF INVENTION: TARGET GENES  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/265,315  
; FILING DATE: March 9, 1999  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/714,918  
; FILING DATE: September 13, 1996  
; APPLICATION NUMBER: 60/009,102  
; FILING DATE: December 22, 1995  
; APPLICATION NUMBER: 60/003,798  
; FILING DATE: September 15, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 240/247  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 513 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-265-315-49

Query Match 73.3%; Score 15.4; DB 4; Length 513;

Best Local Similarity 88.9%; Pred. No. 1.9e+02;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATCATGCTTTTCTCCTG 19

||||| ||||| |||||

Db 386 ATCATACTTTTCTCCTG 403

RESULT 8

US-09-265-315-49

Sequence 49, Application US/09265315

Patent No. 6187541

GENERAL INFORMATION:

APPLICANT: Benton, Bret

APPLICANT: Lee, Ving J.

APPLICANT: Malouin, Francois

APPLICANT: Martin, Patrick K.

APPLICANT: Schmid, Molly B.

APPLICANT: Sun, Dongxu

TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS

TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon &amp; Lyon

STREET: 633 West Fifth Street

STREET: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/265.315

FILING DATE: March 9, 1999

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/714,918

FILING DATE: September 13, 1996

APPLICATION NUMBER: 60/009,102

FILING DATE: December 22, 1995

APPLICATION NUMBER: 60/003,798

FILING DATE: September 15, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 240/247

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 513 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-265-315-49

Query Match 73.3%; Score 15.4; DB 4; Length 513;

Best Local Similarity 88.9%; Pred. No. 1.9e+02;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATCATGCTTTTCTCCTG 19

||||| ||||| |||||

Db 386 ATCATACTTTTCTCCTG 403

RESULT 9

US-09-266-417-49

Sequence 49, Application US/09266417

Patent No. 6228586

GENERAL INFORMATION:

APPLICANT: Benton, Bret

APPLICANT: Lee, Ving J.

APPLICANT: Malouin, Francois

APPLICANT: Martin, Patrick K.

APPLICANT: Schmid, Molly B.

APPLICANT: Sun, Dongxu

TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS

TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon &amp; Lyon

STREET: 633 West Fifth Street

STREET: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/266,417

FILING DATE: March 9, 1999

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/714,918

FILING DATE: September 13, 1996

APPLICATION NUMBER: 60/009,102

FILING DATE: December 22, 1995

APPLICATION NUMBER: 60/003,798

FILING DATE: September 15, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 240/248

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 513 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-266-417-49

Query Match 73.3%; Score 15.4; DB 4; Length 513;

Best Local Similarity 88.9%; Pred. No. 1.9e+02;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATCATGCTTTTCTCCTG 19

||||| ||||| |||||

Db 386 ATCATACTTTTCTCCTG 403

RESULT 9

US-09-265-315-49

Sequence 49, Application US/09265315

Patent No. 6187541

GENERAL INFORMATION:

APPLICANT: Benton, Bret

APPLICANT: Lee, Ving J.

APPLICANT: Malouin, Francois

APPLICANT: Martin, Patrick K.

APPLICANT: Schmid, Molly B.

APPLICANT: Sun, Dongxu

TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS

TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon &amp; Lyon

STREET: 633 West Fifth Street

STREET: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/265,315

FILING DATE: March 9, 1999

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/714,918

FILING DATE: September 13, 1996

APPLICATION NUMBER: 60/009,102

FILING DATE: December 22, 1995

APPLICATION NUMBER: 60/003,798

FILING DATE: September 15, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 240/247

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 513 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-265-315-49

Query Match 73.3%; Score 15.4; DB 4; Length 513;

Best Local Similarity 88.9%; Pred. No. 1.9e+02;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATCATGCTTTTCTCCTG 19

||||| ||||| |||||

Db 386 ATCATACTTTTCTCCTG 403

RESULT 9

US-09-266-417-49

Sequence 49, Application US/09266417

Patent No. 6228586

GENERAL INFORMATION:

APPLICANT: Benton, Bret

APPLICANT: Lee, Ving J.

APPLICANT: Malouin, Francois

APPLICANT: Martin, Patrick K.

APPLICANT: Schmid, Molly B.

APPLICANT: Sun, Dongxu

TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS

TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon &amp; Lyon

STREET: 633 West Fifth Street

STREET: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/266,417

FILING DATE: March 9, 1999

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/714,918

FILING DATE: September 13, 1996

APPLICATION NUMBER: 60/009,102

FILING DATE: December 22, 1995

APPLICATION NUMBER: 60/003,798

FILING DATE: September 15, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 240/248

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 513 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-266-417-49

Query Match 73.3%; Score 15.4; DB 4; Length 513;

Best Local Similarity 88.9%; Pred. No. 1.9e+02;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATCATGCTTTTCTCCTG 19

||||| ||||| |||||

Db 386 ATCATACTTTTCTCCTG 403

Db 386 ATCATACTTTTCTCTCG 403

RESULT 10

US-09-370-838-103/C

Sequence 103, Application US/09370838

Patent No. 6444425

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Lodes, Michael J.

APPLICANT: Mohamath, Roadoh

APPLICANT: Secrist, Heather

TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF

TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE

FILE REFERENCE: 210121.475C1

CURRENT APPLICATION NUMBER: US/09/370,838

CURRENT FILING DATE: 1999-08-09

EARLIER APPLICATION NUMBER: US 09/285,323

EARLIER FILING DATE: 1999-04-02

NUMBER OF SEQ ID NOS: 289

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 103

LENGTH: 741

TYPE: DNA

ORGANISM: Homo sapien

US-09-370-838-103

Query Match 73.3%; Score 15.4; DB 4; Length 741;

Best Local Similarity 94.1%; Pred. No. 1.9e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ATGCTTTTCTCTCTGGG 21

|||||

Db 370 AGCCTTTTCTCTCTGGG 354

RESULT 11

US-08-969-046-15

Sequence 15, Application US/08969046B

Patent No. 6455762

GENERAL INFORMATION:

APPLICANT: Chiang, Vincent Lee C.

APPLICANT: Tsai, Chung-Jui

APPLICANT: Hu, Wen-Jing

TITLE OF INVENTION: Genetic engineering of trees through

TITLE OF INVENTION: manipulation of lignin biosynthesis

FILE REFERENCE: 881.0030S1

CURRENT APPLICATION NUMBER: US/08/969,046B

CURRENT FILING DATE: 1997-11-12

NUMBER OF SEQ ID NOS: 15

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 15

LENGTH: 1927

TYPE: DNA

ORGANISM: Petroselinum crispum

FEATURE:

NAME/KEY: CDS

LOCATION: (64)...(1698)

US-08-969-046-15

Query Match 73.3%; Score 15.4; DB 4; Length 1927;

Best Local Similarity 94.1%; Pred. No. 2.2e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCC 17

|||||

Db 303 CATCATGCTTTTCTCTCC 319

RESULT 12

US-09-052-469-1

Sequence 1, Application US/09052469

Patent No. 6380360

GENERAL INFORMATION:

APPLICANT: Harris et al.

TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE 1 GENE

TITLE OF INVENTION: AND USES THEREOF

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Witcoff, Ltd.

STREET: One Financial Center

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk, 3.50 inch

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/052,469

FILING DATE: Concurrently herewith

PRIOR APPLICATION DATA: 08/422,582

APPLICATION NUMBER: 14-April-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9507766.5

FILING DATE: 13-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9411900.5

FILING DATE: 14-JUN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB PCT/GB94/02822

FILING DATE: 23-DEC-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9326470.3

FILING DATE: 24-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: Williams, Ph.D., Kathleen M.

REGISTRATION NUMBER: 34,380

REFERENCE/DOCKET NUMBER: 3265/74165

TELEPHONE: (617) 345-9100

TELEFAX: (617) 345-9111

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 5631 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: CDNA

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: 1...4842

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 1...5631

OTHER INFORMATION: /function= "Original 3' end of the

OTHER INFORMATION: PKD1 gene"

US-09-052-469-1

Query Match 73.3%; Score 15.4; DB 4; Length 5631;

Best Local Similarity 94.1%; Pred. No. 2.4e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CATGCTTTTCTCTCTGG 20

|||||

Db 2724 CATGCTTTTCTCTGG 2740

RESULT 13

US-08-460-751-1

Sequence 1, Application US/08460751

Patent No. 5891628  
GENERAL INFORMATION:  
APPLICANT: Reeders, Stephen  
APPLICANT: Schneider, Michael  
APPLICANT: Glucksmann, Sandra  
TITLE OF INVENTION: IDENTIFICATION OF POLYCYSTIC KIDNEY  
DISEASE GENE, DIAGNOSTICS AND TREATMENT  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,751  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/413,580  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7638-005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 869-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 1:  
LENGTH: 12912 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..12912  
US-08-460-751-1

Query Match 73.3%; Score 15,4; DB 2; Length 12912;  
Best Local Similarity 94.1%; Pred. No. 2.7e-02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CATGCTTTTCTCCTGG 20  
|||||  
Db 11031 CATGCTTTTCTCCTGG 11047

RESULT 14  
US-09-052-469-5  
Sequence 5, Application US/09052469  
Patent No. 6380360  
GENERAL INFORMATION:  
APPLICANT: Harris et al.  
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE 1 GENE  
AND USES THEREOF  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Witcoff, Ltd.  
STREET: One Financial Center  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02111  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk, 3.50 inch  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
FILING DATE: US/09/052,469  
APPLICATION NUMBER: US/09/052,469  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/422,582  
FILING DATE: 14-April-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9507766.5  
FILING DATE: 13-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9411900.5  
FILING DATE: 14-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB PCT/GB94/02822  
FILING DATE: 23-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9326470.3  
FILING DATE: 24-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams, Ph.D., Kathleen M.  
REGISTRATION NUMBER: 34,380  
REFERENCE/DOCKET NUMBER: 3265/74165  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 345-9100  
TELEFAX: (617) 345-9111  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13807 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2..13018  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 7295..8184  
OTHER INFORMATION: /function= "g alpha 22 fragment"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 6422..7294  
OTHER INFORMATION: /function= "GAP GAMMA PETER  
OTHER INFORMATION: fragment"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 3697..6421  
OTHER INFORMATION: /function= "JH8 fragment"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1373..1701  
OTHER INFORMATION: /function= "S3/S4 PETER fragment"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 2176..2962  
OTHER INFORMATION: /function= "S3/S4 CON2 PETER  
OTHER INFORMATION: fragment"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 2963..3696  
OTHER INFORMATION: /function= "S1/S3 PETER fragment"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 118..1372  
OTHER INFORMATION: /function= "S4/JH13 fragment"  
FEATURE:  
NAME/KEY: misc\_feature

LOCATION: 1..85 /function= "5' COMPLETE [Split]  
OTHER INFORMATION: fragment"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 87..3696 /function= "5' COMPLETE [Split]  
OTHER INFORMATION: fragment"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..85 /function= "6 (5) R cDNA [Split]  
OTHER INFORMATION: fragment"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 87..117 /product= "6 (5) R cDNA [Split]  
OTHER INFORMATION: fragment"  
US-09-052-469-5

Query Match 73.3%; Score 15.4; DB 4; Length 13807;  
Best Local Similarity 94.1%; Pred. No. 2.7e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CATGCTTTTCTCTCTGG 20  
|||||  
Db 10900 CATGCTTTTCTCTGG 10916

RESULT 15  
US-08-658-136-4  
Sequence 4, Application US/08658136  
Patent No. 6071717  
GENERAL INFORMATION:  
APPLICANT: KLINGER, KATHERINE W  
APPLICANT: LANDES, GREGORY M  
APPLICANT: BURN, TIMOTHY C  
APPLICANT: CONNORS, TIMOTHY D  
APPLICANT: DACKOWSKI, WILLIAM  
APPLICANT: GERMINO, GREGORY  
APPLICANT: QIAN, FENG  
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENZYME CORPORATION  
STREET: ONE MOUNTAIN ROAD  
CITY: FRAMINGHAM  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 01701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/658,136  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: LASSEN, ELIZABETH  
REGISTRATION NUMBER: 31,845  
REFERENCE/DOCKET NUMBER: GEN4-17.8  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 508-872-8400  
TELEFAX: 508-872-5415  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14060 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA

FEATURE:  
NAME/KEY: CDS  
LOCATION: 135..13040  
US-08-658-136-4  
Query Match 73.3%; Score 15.4; DB 3; Length 14060;  
Best Local Similarity 94.1%; Pred. No. 2.7e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 4 CATGCTTTTCTCTCTGG 20  
|||||  
Db 11162 CATGCTTTTCTCTGG 11178  
Search completed: July 8, 2003, 09:32:17  
Job time : 30.0402 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 19:09:54 ; Search time 121.288 seconds  
(without alignments)  
273.390 Million cell updates/sec.

Title: US-09-647-780A-17

Perfect score: 21

Sequence: 1 catcatgctttttctctggg 21

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1105431 seqs, 789497651 residues

Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications-NA:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17.4	82.9	329	9	US-10-095-672A-3
2	17.4	82.9	1297	10	US-09-263-959-288
3	17.4	82.9	684973	10	US-09-263-959-1
4	16.8	80.0	481	9	US-09-918-995-24241
5	16.8	80.0	484	10	US-09-998-598-41
6	16.8	80.0	566	10	US-09-998-598-70
7	16.8	80.0	1100	10	US-09-935-390A-13
8	16.8	80.0	1389	10	US-09-998-598-362
9	16.8	80.0	1425	9	US-09-746-783-131
10	16.8	80.0	1480	9	US-10-023-282-207
11	16.4	78.1	482	10	US-09-881-823-19
12	16.4	78.1	1632	10	US-09-917-800A-1522
13	16.4	78.1	2000	9	US-09-938-842A-3679
14	16.2	77.1	1156	10	US-09-778-844-54
15	16.2	77.1	1160	10	US-09-778-844-55
16	16.2	77.1	536165	9	US-09-939-964-1
17	16	76.2	489	9	US-10-125-540-124
18	16	76.2	489	10	US-09-764-870-124
19	16	76.2	1038	9	US-09-864-921-11

20	16	76.2	1107	9	US-09-864-921-84	Sequence 84, Appl
21	16	76.2	1245	10	US-09-833-381-1974	Sequence 1974, Ap
22	16	76.2	3111	9	US-10-295-981-56	Sequence 56, Appl
23	16	76.2	3111	10	US-09-728-721-56	Sequence 54, Appl
24	16	76.2	4244	9	US-10-295-981-54	Sequence 54, Appl
25	16	76.2	4244	10	US-09-728-721-54	Sequence 79, Appl
26	15.8	75.2	226	9	US-09-907-969-79	Sequence 79, Appl
27	15.8	75.2	226	9	US-09-907-969-253	Sequence 253, App
28	15.8	75.2	226	9	US-09-907-969-254	Sequence 254, App
29	15.8	75.2	226	9	US-10-198-053-79	Sequence 79, Appl
30	15.8	75.2	226	9	US-10-198-053-253	Sequence 253, App
31	15.8	75.2	226	9	US-10-198-053-254	Sequence 254, App
32	15.8	75.2	226	10	US-09-884-441-79	Sequence 79, Appl
33	15.8	75.2	226	10	US-09-884-441-253	Sequence 253, App
34	15.8	75.2	226	10	US-09-884-441-254	Sequence 254, App
35	15.8	75.2	361	10	US-09-834-975-648	Sequence 648, App
36	15.8	75.2	419	10	US-09-983-965-5635	Sequence 5635, Ap
37	15.8	75.2	432	10	US-09-867-701-2945	Sequence 2945, Ap
38	15.8	75.2	457	10	US-09-867-701-6165	Sequence 6165, Ap
39	15.8	75.2	493	10	US-09-833-740-6	Sequence 6, Appl
40	15.8	75.2	581	10	US-09-834-975-314	Sequence 314, App
41	15.8	75.2	981	10	US-09-822-849A-204	Sequence 204, App
42	15.8	75.2	1812	9	US-10-128-714-2487	Sequence 2487, Ap
43	15.8	75.2	1812	9	US-10-128-714-7487	Sequence 7487, Ap
44	15.8	75.2	1974	9	US-10-128-714-1487	Sequence 1487, Ap
45	15.8	75.2	1974	9	US-10-128-714-6487	Sequence 6487, Ap

#### ALIGNMENTS

RESULT 1  
US-10-095-672A-3  
Sequence 3, Application US/10095672A  
Publication No. US20030068628A1  
GENERAL INFORMATION:  
APPLICANT: Hercend, Thierry  
APPLICANT: Triebel, Frederic  
APPLICANT: Roman-Roman, Sergio  
APPLICANT: Ferradini, Laurent  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR VARIABLE REGIONS OF BETA CHAINS OF HUMAN T LYMPHOCYTE RECEPTORS, CORRESPONDING  
TITLE OF INVENTION: PEPTIDE SEGMENTS AND THE DIAGNOSTIC AND THERAPEUTIC USES  
FILE REFERENCE: 146.1158-CON-DIV-2-CON  
CURRENT APPLICATION NUMBER: US/10/095,672A  
CURRENT FILING DATE: 2002-03-12  
PRIOR APPLICATION NUMBER: US 08/437,353  
PRIOR FILING DATE: 1995-05-09  
PRIOR APPLICATION NUMBER: US 08/423,383  
PRIOR FILING DATE: 1995-04-14  
PRIOR APPLICATION NUMBER: US 07/934,530  
PRIOR FILING DATE: 1992-11-23  
PRIOR APPLICATION NUMBER: PCT/FR92/00130  
PRIOR FILING DATE: 1992-02-23  
PRIOR APPLICATION NUMBER: FR/91/01613  
PRIOR FILING DATE: 1990-02-12  
PRIOR APPLICATION NUMBER: FR/91/04523  
PRIOR FILING DATE: 1991-02-12  
NUMBER OF SEQ ID NOS: 87  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 329  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: IGR b 04; V BETA w23  
US-10-095-672A-3

Query Match 82.9% Score 17.4; DB 9; Length 329;  
Best Local Similarity 94.7%; Pred. No. 1.2e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCATGCTTTTCTCTCTGGG 21  
Db 16 TCATGCTTTGTCTCTCTGGG 34

## RESULT 2

US-09-263-959-288  
; Sequence 288, Application US/09263959  
; Patent No. US20020150891A1  
; GENERAL INFORMATION:  
; APPLICANT: Hood, Leroy E.  
; APPLICANT: Rowen, Lee  
; APPLICANT: Koop, Ben F.  
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTI  
; NUMBER OF SEQUENCES: 1279  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 05-MAR-1999  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcmasters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 920010.426C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031

## INFORMATION FOR SEQ ID NO:

288:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1297 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-263-959-288

Query Match 82.9%; Score 17.4; DB 10; Length 1297;  
Best Local Similarity 94.7%; Pred. No. 1.3e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCATGCTTTTCTCTCTGGG 21  
Db 678 TCATGCTTTGTCTCTCTGGG 696

## RESULT 3

US-09-263-959-1  
; Sequence 1, Application US/09263959  
; Patent No. US20020150891A1  
; GENERAL INFORMATION:  
; APPLICANT: Hood, Leroy E.  
; APPLICANT: Rowen, Lee  
; APPLICANT: Koop, Ben F.  
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTI  
; NUMBER OF SEQUENCES: 1279  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 05-MAR-1999  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcmasters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 920010.426C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 684973 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-263-959-1

Query Match 82.9%; Score 17.4; DB 10; Length 684973;  
Best Local Similarity 94.7%; Pred. No. 2.3e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCATGCTTTTCTCTCTGGG 21  
Db 370711 TCATGCTTTGTCTCTCTGGG 370729

## RESULT 4

US-09-918-995-24241.  
; Sequence 24241, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 24241  
; LENGTH: 481  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(481)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-24241

Query Match 80.0%; Score 16.8; DB 9; Length 481;  
Best Local Similarity 90.0%; Pred. No. 2.3e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCTGG 20  
Db 385 CATCTGCTTTGTCTCTGG 404

## RESULT 5

US-09-998-598-41  
; Sequence 41, Application US/09998598  
; Patent No. US20020150922A1  
; GENERAL INFORMATION:  
; APPLICANT: Stolk, John A.  
; APPLICANT: Xu, Jiangchun



Publication No. US20030044935A1  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
McCoy, John M.  
LaValle, Edward R.  
Racie, Lisa A.  
Treacy, Maurice  
Spaulding, Vikki  
Agostino, Michael J.  
Howes, Steven H.  
Fechtel, Kim  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM  
NUMBER OF SEQUENCES: 231  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/746,783  
FILING DATE: 21-Dec-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Milasincic, Debra J.  
REGISTRATION NUMBER: 46,931  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 131:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1425 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "oligonucleotide"  
SEQUENCE DESCRIPTION: SEQ ID NO: 131:  
US-09-746-783-131  
Query Match 80.08; Score 16.8; DB 9; Length 1425;  
Best Local Similarity 90.08; Pred. No. 2.5e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 CATCATGCTTTTCTCTCTGG 20  
|||||  
Db 318 CATCTGCTTTGTCTCTCTGG 337  
RESULT 10  
US-10-023-282-207  
Sequence 207, Application US/10023282  
Publication No. US20030092893A1  
GENERAL INFORMATION:  
APPLICANT: Young et al.  
TITLE OF INVENTION: 207 Human Secreted Proteins  
FILE REFERENCE: P2007P1  
CURRENT APPLICATION NUMBER: US/10/023,282  
CURRENT FILING DATE: 2001-12-20  
EARLIER APPLICATION NUMBER: 09/205,258  
EARLIER FILING DATE: 1998-12-04  
EARLIER APPLICATION NUMBER: PCT/US98/11422  
EARLIER FILING DATE: 1998-06-04  
EARLIER APPLICATION NUMBER: 60/048,885  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,375  
EARLIER FILING DATE: 1998-07-15  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,881  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,880  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,896  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,020  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,876  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,895  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,884  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,894  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,971  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,882  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,899  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,893  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,900  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,901  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,892  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,915  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,019  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,972  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,916  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,373  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,875  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,374  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,917  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,949  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15

EARLIER APPLICATION NUMBER: 60/094,657;  
EARLIER FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 207  
LENGTH: 1480  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-023-282-207

Query Match 80.0%; Score 16.8; DB 9; Length 1480;  
Best Local Similarity 90.0%; Pred. No. 2.5e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCTGG 20  
||||| ||||| ||||| |||||  
DB 379 CATCTGCTTTTCTCTCTGG 398

RESULT 11  
US-09-881-823-19  
Sequence 19, Application US/09881823  
Patent No. US20020068068A1  
GENERAL INFORMATION:  
APPLICANT: SHI, WENYUAN  
APPLICANT: ANDERSON, MAXWELL  
APPLICANT: MORRISON, SHERIE  
APPLICANT: TRINH, RYAN  
APPLICANT: WIMS, LETITIA  
APPLICANT: CHEN, LI  
TITLE OF INVENTION: Method for the Treatment and Prevention of Dental Caries  
FILE REFERENCE: 22851-032  
CURRENT APPLICATION NUMBER: US/09/881,823  
CURRENT FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 07/378,577  
PRIOR FILING DATE: 1999-08-20  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 19  
LENGTH: 482  
TYPE: DNA  
ORGANISM: Murine  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (27)...(482)  
US-09-881-823-19

Query Match 78.1%; Score 16.4; DB 10; Length 482;  
Best Local Similarity 94.4%; Pred. No. 3.5e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCATGCTTTTCTCTCTGG 20  
||||| ||||| ||||| |||||  
DB 43 TCATGCTTTTCTCTCTGG 60

RESULT 12  
US-09-917-800A-1522  
Sequence 1522, Application US/09917800A  
Patent No. US20020119462A1  
GENERAL INFORMATION:  
APPLICANT: Mendrick, Donna  
APPLICANT: Porter, Mark  
APPLICANT: Johnson, Kory  
APPLICANT: Castle, Arthur  
APPLICANT: Elashoff, Michael  
APPLICANT: Gene Logic, Inc.  
TITLE OF INVENTION: Molecular Toxicology Modeling  
FILE REFERENCE: 44921-5038-US  
CURRENT APPLICATION NUMBER: US/09/917,800A  
CURRENT FILING DATE: 2001-07-31  
PRIOR APPLICATION NUMBER: US 60/222,040  
PRIOR FILING DATE: 2000-07-31

PRIOR APPLICATION NUMBER: US 60/222,880  
PRIOR FILING DATE: 2000-11-02  
PRIOR APPLICATION NUMBER: US 60/290,029  
PRIOR FILING DATE: 2001-05-11  
PRIOR APPLICATION NUMBER: US 60/290,645  
PRIOR FILING DATE: 2001-05-15  
PRIOR APPLICATION NUMBER: US 60/292,336  
PRIOR FILING DATE: 2001-05-22  
PRIOR APPLICATION NUMBER: US 60/295,798  
PRIOR FILING DATE: 2001-06-06  
PRIOR APPLICATION NUMBER: US 60/297,457  
PRIOR FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/298,884  
PRIOR FILING DATE: 2001-06-19  
PRIOR APPLICATION NUMBER: US 60/303,459  
PRIOR FILING DATE: 2001-07-09  
NUMBER OF SEQ ID NOS: 1740  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1522  
LENGTH: 1632  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. US20020119462A1 X98517  
US-09-917-800A-1522

Query Match 78.1%; Score 16.4; DB 10; Length 1632;  
Best Local Similarity 94.4%; Pred. No. 3.9e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CATGCTTTTCTCTCTGG 21  
||||| ||||| ||||| |||||  
DB 579 CATGCTTTTCTCTCTGG 596

RESULT 13  
US-09-938-842A-3679/c  
Sequence 3679, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kreps, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
FILE REFERENCE: SCRIPI300-3  
CURRENT APPLICATION NUMBER: US/09/938,842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227,866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264,647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300,111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 3679  
LENGTH: 2000  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-3679

Query Match 78.1%; Score 16.4; DB 9; Length 2000;  
Best Local Similarity 94.4%; Pred. No. 4e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCT 18  
||||| ||||| ||||| |||||  
DB 1812 CATATGCTTTTCTCTCT 1795

RESULT 14  
US-09-778-844-54/c

; Sequence 54, Application US/09778844  
; Patent No. US20020150971A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHANSEN, JEANETTE ELISABETH  
; APPLICANT: SCHALLING, MARTIN  
; TITLE OF INVENTION: NUCLEIC ACIDS AND POLYPEPTIDES FOR CONTROLLING FOOD  
; TITLE OF INVENTION: INTAKE AND/OR BODY WEIGHT  
; FILE REFERENCE: 030307/0195  
; CURRENT APPLICATION NUMBER: US/09/778,844  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 54  
; LENGTH: 1156  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: EMBL No. US20020150971A1 q9ulg1  
US-09-778-844-54

Query Match 77.1%; Score 16.2; DB 10; Length 1156;  
Best Local Similarity 85.7%; Pred. No. 4.6e+02;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCTGGG 21  
||||| ||||||| |||||  
Db 516 CATCTCCCTTTTCTCTCTGGG 496

RESULT 15  
US-09-778-844-55/c  
; Sequence 55, Application US/09778844  
; Patent No. US20020150971A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHANSEN, JEANETTE ELISABETH  
; APPLICANT: SCHALLING, MARTIN  
; TITLE OF INVENTION: NUCLEIC ACIDS AND POLYPEPTIDES FOR CONTROLLING FOOD  
; TITLE OF INVENTION: INTAKE AND/OR BODY WEIGHT  
; FILE REFERENCE: 030307/0195  
; CURRENT APPLICATION NUMBER: US/09/778,844  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 55  
; LENGTH: 1160  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: EMBL No. US20020150971A1 q9ulg1  
US-09-778-844-55

Query Match 77.1%; Score 16.2; DB 10; Length 1160;  
Best Local Similarity 85.7%; Pred. No. 4.6e+02;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCTGGG 21  
||||| ||||||| |||||  
Db 117 CATCTCCCTTTTCTCTCTGGG 97

Search completed: July 9, 2003, 02:22:32  
Job time : 123.338 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 00:47:28 ; Search time 1064.65 Seconds  
(without alignments)  
319.453 Million cell updates/sec

Title: US-09-647-780A-17  
Perfect score: 21  
Sequence: 1 catcatgcttttctctctggg 21

Scoring table: IDENTITY\_NUC  
Gapop:10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

```
EST:**
1: em_estba:**
2: em_esthum:**
3: em_estin:**
4: em_estnu:**
5: em_estov:**
6: em_estpl:**
7: em_estro:**
8: em_htc:**
9: gb_est1:**
10: gb_est2:**
11: gb_htc:**
12: gb_est3:**
13: gb_est4:**
14: gb_est5:**
15: em_estfun:**
16: em_estom:**
17: gb_gss:**
18: em_gss_hum:**
19: em_gss_inv:**
20: em_gss_pin:**
21: em_gss_vrt:**
22: em_gss_fun:**
23: em_gss_mam:**
24: em_gss_mus:**
25: em_gss_other:**
26: em_gss_pro:**
27: em_gss_rod:**
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	18.4	87.6	256	13	BM494943
c 2	18.4	87.6	390	10	AW521249
c 3	18.4	87.6	397	9	AI872268
c 4	18	85.7	291	13	BI435848
c 5	18	85.7	331	14	BQ120752
c 6	18	85.7	336	13	BM108635

c 7	18	85.7	378	9	AJ489153
c 8	18	85.7	416	13	BM407629
c 9	18	85.7	463	12	BG019241
c 10	18	85.7	504	13	BM109913
c 11	18	85.7	504	13	BM405275
c 12	18	85.7	518	12	BG890211
c 13	18	85.7	519	13	BM407419
c 14	18	85.7	535	13	BM406937
c 15	18	85.7	537	13	BM407841
c 16	18	85.7	562	10	AW637123
c 17	18	85.7	571	14	BQ120751
c 18	18	85.7	693	17	AQ050592
c 19	18	85.7	929	9	AL580841
c 20	17.8	84.8	352	14	T78565
c 21	17.8	84.8	371	14	N28602
c 22	17.8	84.8	378	10	BE680289
c 23	17.8	84.8	432	17	A2423292
c 24	17.8	84.8	451	13	BI450122
c 25	17.8	84.8	472	12	BF613826
c 26	17.8	84.8	517	10	BE650548
c 27	17.8	84.8	562	17	AZ896498
c 28	17.8	84.8	650	9	AL652380
c 29	17.8	84.8	658	17	AQ585863
c 30	17.8	84.8	686	17	AZ285976
c 31	17.8	84.8	1112	12	BE902285
c 32	17.8	84.8	1176	14	BQ071808
c 33	17.4	82.9	180	9	AL369698
c 34	17.4	82.9	203	9	AA329883
c 35	17.4	82.9	246	12	BF773299
c 36	17.4	82.9	297	12	BG448196
c 37	17.4	82.9	315	13	BI268295
c 38	17.4	82.9	357	12	BF946678
c 39	17.4	82.9	422	9	AJ502366
c 40	17.4	82.9	422	9	AJ502929
c 41	17.4	82.9	433	17	AZ298482
c 42	17.4	82.9	445	13	BM208993
c 43	17.4	82.9	459	9	AA315638
c 44	17.4	82.9	473	9	AI177440
c 45	17.4	82.9	477	12	BG802076

#### ALIGNMENTS

RESULT 1  
BM494943/c  
LOCUS  
DEFINITION  
256 bp. mRNA linear EST 11-FEB-2002  
IPCGBr1\_6\_B10\_21 Ictalurus punctatus Brainl primary library  
Ictalurus punctatus cdna clone IPCGBr1\_6\_B10\_21\_15Feb00\_077 5',  
mRNA sequence.  
BM494943  
EST.  
BM494943.1 GI:18646124  
channel catfish.  
Ictalurus punctatus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;  
Ictaluridae; Ictalurus.  
1 (bases 1 to 256).  
Noneman,D.J. and Waldbieser,G.C.  
Characterization of a brain cdna library from adult channel catfish  
(Ictalurus punctatus)  
Unpublished (2002)  
JOURNAL  
COMMENT  
Contact: Waldbieser GC  
Catfish Genetics Research Unit  
USDA-Agricultural Research Service  
141 Experiment Station Road, Stoneville, MS 38776, USA  
Tel: 662 686 3593  
Fax: 662 686 3567  
Email: gwaldbieser@ars.usda.gov  
Single pass sequencing. Bases called with Phred v0.000925.c.c. Low  
quality bases and vector trimmed with Lucy v1.16.  
Plate: Br1\_6 row: B column: 10

Seq primer: M13 Reverse.  
 FEATURES  
 source

Location/Qualifiers  
 1..256  
 /organism="Ictalurus punctatus"  
 /strain="USDA103"  
 /db\_xref="taxon:7998"  
 /clone="ipGB1\_6\_B10\_21\_15Feb00\_077"  
 /clone\_lib="Ictalurus punctatus Brain1 primary library"  
 /sex="female and male"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pSport1; Site\_1: Sal I;  
 Site\_2: Not I; Primary library"  
 BASE COUNT 81 a 52 c 81 g 42 t  
 ORIGIN

Query Match 87.6%; Score 18.4; DB 13; Length 256;  
 Best Local Similarity 95.0%; Pred. NO. 7.8e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCGG 20  
 |||||||  
 Db 136 CATCATGCTTTTCTCTTG 117

RESULT 2  
 AW521249  
 LOCUS 390 bp mRNA linear EST 06-MAR-2000  
 DEFINITION UI-R-800-agk-h-12-0-UI.s1 UI-R-800 Rattus norvegicus cDNA clone  
 UI-R-800-agk-h-12-0-UI 3', mRNA sequence.  
 ACCESSION AW521249  
 VERSION AW521249.1 GI:7163627  
 KEYWORDS EST.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 390)  
 Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 97044477  
 Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.wesg.uiowa.edu

The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. The sequence tag present in the cDNA between the NotI site  
 and the oligo-dT track served to identify it as a clone from the  
 normalized medulla library cDNA library Preparation: M.B. Soares  
 Lab Clone distribution: clones will be available through Research  
 Genetics (www.resgen.com)  
 Seq primer: M13 Forward  
 POLYA=Yes.  
 FEATURES  
 source

Location/Qualifiers  
 1..390  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-800-agk-h-12-0-UI"  
 /clone\_lib="UI-R-800"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: p773D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; The library  
 (UI-R-800) is a subtracted library derived from a mixture

of the following tissues: thalamus, cerebellum,  
 hypothalamus, medulla, pons, midbrain, cerebral  
 corpus striatum and hippocampus. For a detailed  
 description of the library from which this clone was  
 derived, please visit our web site at  
 ratest.eng.uiowa.edu. The subtraction has been previously  
 described in (Bonaldo, Lennon and Soares, Genome Research  
 6:791-806, 1996)  
 TAG\_LIB=UI-R-800  
 TAG\_TISSUE=medulla  
 TAG\_SEQ=GAACCG  
 BASE COUNT 134 a 61 g 132 t  
 ORIGIN

Query Match 87.6%; Score 18.4; DB 10; Length 390;  
 Best Local Similarity 95.0%; Pred. NO. 9e+02; 1; Indels 0; Gaps 0;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCGG 20  
 |||||||  
 Db 210 CATCATGCTTTTCTCTCGG 229

RESULT 3  
 AI872268

LOCUS 397 bp mRNA linear EST 07-MAR-2000  
 DEFINITION tz62h02.x1 NCI\_CGAP\_Ov35 Homo sapiens cDNA clone IMAGE:2293203 3'  
 similar to contains Alu repetitive element;; mRNA sequence.  
 ACCESSION AI872268  
 VERSION AI872268.1 GI:5546317  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 397)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael.  
 R. Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bbrp/image/image.html  
 Insert Length: 1416 Std Error: 0.00  
 Seq primer: -400P from GIBCO.

FEATURES  
 source

Location/Qualifiers  
 1..397  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2293203"  
 /clone\_lib="NCI\_CGAP\_Ov35"  
 /tissue\_type="tumor, 5 pooled (see description)"  
 /lab\_host="DH10B"  
 /note="Organ: ovary; Vector: pCMV-SPORT6; Site\_1: SalI;  
 Site\_2: NotI; This library represents the normalized  
 version of NCI\_CGAP\_Ov23. Cloned unidirectionally.  
 Primer: Oligo dT. Average insert size 0.86 kb. Tumor  
 types include: mixed Mullerian tumor, papillary serous,  
 clear cell, spindle cell. All are primary tumors,  
 metastasis positive. Constructed by Life Technologies."  
 BASE COUNT 103 a 100 c 86 g 108 t  
 ORIGIN

Query Match 87.6%; Score 18.4; DB 9; Length 397;  
 Best Local Similarity 95.0%; Pred. No. 9.1e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;





```

/clone="cpro1h16"
/clone_lib="potato roots"
/tissue_type="roots"
/dev_stage="in vitro grown stem cuttings"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University, Tanksley lab;
sequencing: The Institute for Genomic Research. Roots were
isolated from in vitro grown stem cuttings on CM medium.
Roots were isolated two weeks after placing the stem
cuttings from in vitro grown plants on medium."
BASE COUNT      149 a      95 g      46 t
ORIGIN
Query Match      85.7%; Score 18; DB 13; Length 336;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATCATGCTTTTCTCCT 18
|||||
Db 271 CATCATGCTTTTCTCCT 254

RESULT 7
AJ489153/c
AJ489153/LOCUS
DEFINITION AJ489153 Solanum tuberosum cv. Saturna linear EST 30-JUL-2002
S3b7, mRNA sequence.
ACCESSION AJ489153
VERSION AJ489153.1 GI:22022271
KEYWORDS
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 378)
AUTHORS Gebhardt, C., Walkemeier, B., Henselewski, H., Barakat, A., Delseny, M.
and Stueber, K.
TITLE Comparative mapping between potato (Solanum tuberosum) and
Arabidopsis thaliana reveals structurally conserved domains and
ancient duplications in the potato genome
JOURNAL unpublished (2002)
COMMENT Contact: Gebhardt C
Plant breeding and yield physiology
MPI for Breeding Research
Carl-von-Linne Weg 10, Cologne, D-50829, Germany.
FEATURES
source
Location/Qualifiers
1..378
/organism="Solanum tuberosum"
/cultivar="Saturna"
/db_xref="taxon:4113"
/map="v"
/clone="S3b7"
/clone_lib="Solanum tuberosum cv. Saturna"
BASE COUNT      159 a      57 c      95 g      67 t
ORIGIN
Query Match      85.7%; Score 18; DB 9; Length 378;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATCATGCTTTTCTCCT 18
|||||
Db 274 CATCATGCTTTTCTCCT 257

RESULT 8
BM407629/c
BM407629/LOCUS
DEFINITION ESF581956 potato roots Solanum tuberosum cdna clone cpro31d6 5' end
mRNA sequence.
ACCESSION BM407629

```

```

VERSION BM407629.1 GI:18259259
KEYWORDS EST.
SOURCE Solanum tuberosum
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 416)
AUTHORS van der Hoeven, R., Sun, H., Karanymcheva, S.A., Tsai, J., Van Aken, S.,
Tanksley, S., Chleming, A., Bougri, O., Buell, C.R., Ronning, C.,
Generation of ESTs from potato roots
Unpublished (2001)
TITLE JOURNAL
COMMENT Contact: Research Genetics, Libraries Division
Tel: 1-800-711-6195
Email: cdna@resgen.com
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: T3.
FEATURES
source
Location/Qualifiers
1..416
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cpro31d6"
/clone_lib="potato roots"
/tissue_type="roots"
/dev_stage="in vitro grown stem cuttings"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University, Tanksley lab;
sequencing: The Institute for Genomic Research. Roots were
isolated from in vitro grown stem cuttings on CM medium.
Roots were isolated two weeks after placing the stem
cuttings from in vitro grown plants on medium."
BASE COUNT      174 a      54 c      102 g      76 t
ORIGIN
Query Match      85.7%; Score 18; DB 13; Length 416;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATCATGCTTTTCTCCT 18
|||||
Db 284 CATCATGCTTTTCTCCT 267

RESULT 9
BG019241/c
BG019241/LOCUS
DEFINITION daa76d05.y1 Cho Li treated gastrula Xenopus laevis cdna clone
IMAGE:4061625 5' similar to SW:NA4M_HUMAN O95182 NADH-UBIQUINONE
OXIDOREDUCTASE SUBUNIT B14.5A ;, mRNA sequence.
ACCESSION BG019241
VERSION BG019241.1 GI:12475248
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 463)
AUTHORS Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D.,
Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person,
B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,
Waterston, R. and Willson, R.
WashU Xenopus EST project, 1999
Unpublished (1999)
TITLE JOURNAL
COMMENT Contact: Sandy Clifton, Ph.D.
WashU Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

```

Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Library constructed by C. Hashimoto, Ph.D. in the Laboratory of K. Cho, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov  
Seq primer: -40RP from Gibco  
High quality sequence stop: 400.

# FEATURES

Location/Qualifiers  
1..463

/organism="Xenopus laevis"  
/db\_xref="taxon:8355"  
/clone="IMAGE:4061625"  
/clone\_lib="Cho Li treated gastrula"  
/tissue\_type="gastrula, Li treated"  
/lab\_host="DHL0B"

/note="vector: pBluescript KS+; Site\_1: NotI; Site\_2: EcoRI; 1st strand was primed with a Not I - oligo(GT) primer, double-stranded cDNA was cloned into the Not I and Eco RI sites of pBluescript KS+. Library was constructed by C. Hashimoto, Ph.D., in the laboratory of K. Cho, Ph.D. (Department of Developmental and Cell Biology, University of California, Irvine)."

BASE COUNT 104 a 129 c 116 g 114 t

# ORIGIN

Query Match 85.7%; Score 18; DB 12; Length 463;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCT 18  
|||||  
DB 335 CATCATGCTTTTCTCTCT 318

# RESULT 10

BM109913/c  
LOCUS BM109913 504 bp mRNA linear EST 26-NOV-2001  
DEFINITION EST55749 potato roots Solanum tuberosum cDNA clone cPRO5N21 5' end  
/ mRNA sequence.

ACCESSION BM109913  
VERSION BM109913.1 GI:17071177  
KEYWORDS EST.  
SOURCE potato.

# ORGANISM

Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

1 (bases 1 to 504)

van der Hoeven, R., Sun, H., Karameycheva, S.A., Tsai, J., Van Aken, S., Utterback, T., Chiemiango, A., Bougri, O., Buell, C.R., Ronning, C., Tanksley, S. and Baker, B.

Generation of ESTs from potato roots

Unpublished (2001)

Contact: Research Genetics, Libraries Division

Tel: 1-800-711-6195

Email: cdna@resgen.com

For clone info: please contact Research Genetics, Libraries Division tel 1-800-711-6195, email cdna@resgen.com

Seq primer: T3.

# FEATURES

source

Location/Qualifiers  
1..504

/organism="Solanum tuberosum"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="cPRO5N21"  
/clone\_lib="potato roots"  
/tissue\_type="roots"  
/dev\_stage="in vitro grown stem cuttings"  
/lab\_host="SOLR"

/note="vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; supplier: Cornell University, Tanksley lab;

sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."

BASE COUNT 206 a 73 c 120 g 105 t

# ORIGIN

Query Match 85.7%; Score 18; DB 13; Length 504;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCT 18  
|||||  
DB 285 CATCATGCTTTTCTCTCT 268

# RESULT 11

BM405275/c  
LOCUS BM405275 504 bp mRNA linear EST 22-JAN-2002  
DEFINITION EST579602 potato roots Solanum tuberosum cDNA clone cPRO23K18 5'  
end, mRNA sequence.

ACCESSION BM405275

VERSION BM405275.1 GI:18256785

KEYWORDS EST.

SOURCE potato.

ORGANISM Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

1 (bases 1 to 504)

van der Hoeven, R., Sun, H., Karameycheva, S.A., Tsai, J., Van Aken, S., Utterback, T., Chiemiango, A., Bougri, O., Buell, C.R., Ronning, C., Tanksley, S. and Baker, B.

Generation of ESTs from potato roots

Unpublished (2001)

Contact: Research Genetics, Libraries Division

Tel: 1-800-711-6195

Email: cdna@resgen.com

For clone info: please contact Research Genetics, Libraries Division tel 1-800-711-6195, email cdna@resgen.com

Seq primer: T3.

# FEATURES

source

Location/Qualifiers  
1..504

/organism="Solanum tuberosum"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="cPRO23K18"  
/clone\_lib="potato roots"  
/tissue\_type="roots"  
/dev\_stage="in vitro grown stem cuttings"  
/lab\_host="SOLR"

/note="vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; supplier: Cornell University, Tanksley lab; sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."

BASE COUNT 200 a 68 c 122 g 114 t

# ORIGIN

Query Match 85.7%; Score 18; DB 13; Length 504;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCTCT 18  
|||||  
DB 255 CATCATGCTTTTCTCTCT 238

# RESULT 12

BG890211/c  
LOCUS BG890211 518 bp mRNA linear EST 30-MAY-2001  
DEFINITION EST516062 cSTD Solanum tuberosum cDNA clone cSTD17H15 5' sequence,



Query Match 85.7%; Score 18; DB 13; Length 535;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCCT 18  
 |||||  
 Db 272 CATCATGCTTTTCTCCT 255

RESULT 15  
 BM407841/c  
 LOCUS BM407841 537 bp mRNA linear EST 22-JAN-2002  
 DEFINITION EST582168 potato roots Solanum tuberosum cDNA clone CPRO32M20 5'  
 end, mRNA sequence.  
 ACCESSION BM407841  
 VERSION BM407841.1 GI:18259471  
 KEYWORDS EST.  
 SOURCE potato.  
 ORGANISM Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 REFERENCE 1 (bases 1 to 537)  
 AUTHORS van der Hoeven, R., Sun, H., Karamycheva, S.A., Tsai, J., Van Aken, S.,  
 Utterback, T., Chiemingo, A., Bougri, O., Buell, C.R., Ronning, C.,  
 Tanksley, S. and Baker, B.  
 TITLE Generation of ESTs from potato roots  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Research Genetics, Libraries Division  
 Tel: 1-800-711-6195  
 Email: cdna@resgen.com  
 For clone info: please contact Research Genetics, Libraries  
 Division tel 1-800-711-6195, email cdna@resgen.com  
 Seq primer: T3.

FEATURES  
 source  
 1..537  
 /organism="Solanum tuberosum"  
 /cultivar="kennebec"  
 /db\_xref="taxon:4113"  
 /clone="CPRO32M20"  
 /clone\_lib="potato roots"  
 /tissue\_type="roots"  
 /dev\_stage="in vitro grown stem cuttings"  
 /lab\_host="SOLR"  
 /note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
 XhoI; supplier: Cornell University, Tanksley lab;  
 sequencing: The Institute for Genomic Research. Roots were  
 isolated from in vitro grown stem cuttings on CM medium.  
 Roots were isolated two weeks after placing the stem  
 cuttings from in vitro grown plants on medium."

BASE COUNT 211 a 71 c 126 g 129 t  
 ORIGIN

Query Match 85.7%; Score 18; DB 13; Length 537;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATCATGCTTTTCTCCT 18  
 |||||  
 Db 267 CATCATGCTTTTCTCCT 250

Search completed: July 8, 2003, 09:22:04  
 Job time : 1068.8 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 00:45:53 ; Search time 231.102 seconds  
(without alignments)  
2644.537 Million cell updates/sec

Title: US-09-647-780A-18  
Perfect score: 21  
Sequence: 1 ccgaagtttcttgaggctcc 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.man.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	21	100.0	21	6	AX014718 Sequence
2	21	100.0	2765	6	AX014701 Sequence
3	21	100.0	174953	2	AC094732 Rattus no
4	19.4	92.4	2583	10	AF302075 Mus muscu
5	19.4	92.4	2601	10	AF157106 Mus muscu
6	19.4	92.4	2652	10	AF302076 Mus muscu
7	19.4	92.4	2694	10	AF302077 Mus muscu
8	19.4	92.4	2892	10	AF157105 Mus muscu
9	19.4	92.4	2925	6	AX033272 Sequence
10	19.4	92.4	2925	10	AF176569 Mus muscu
11	19.4	92.4	67142	2	AC120432 Mus muscu
12	19.4	92.4	208249	2	AL607032 Mus muscu
13	18.4	87.6	110000	2	TBCHRIA_06
14	17.8	84.8	2076	6	AX146976 Sequence
15	17.8	84.8	2232	6	AX319864 Sequence
16	17.8	84.8	2262	6	AX146980 Sequence
17	17.8	84.8	2340	6	AX146978 Sequence
18	17.8	84.8	2340	6	AX473102 Sequence
19	17.8	84.8	2636	6	AX139743 Sequence
20	17.8	84.8	2663	6	AX139745 Sequence
21	17.8	84.8	2676	6	AX033274 Sequence
22	17.8	84.8	2714	6	AX139741 Sequence
23	17.8	84.8	2784	9	AF336981 Homo sapi
24	17.8	84.8	2850	9	AK093058 Homo sapi
25	17.8	84.8	2893	6	AX356951 Sequence
26	17.8	84.8	2893	6	AX463057 Sequence
27	17.8	84.8	2953	6	AX473100 Sequence
28	17.8	84.8	2975	6	AX356955 Sequence
29	17.8	84.8	2975	6	AX463058 Sequence
30	17.8	84.8	141277	2	AC129117 Rattus no
31	17.8	84.8	154736	9	AL139246 Human DNA
32	17.8	84.8	160957	9	AC114487 Homo sapi
33	17.8	84.8	181059	2	AC107435 Rattus no
34	17.4	82.9	30838	2	AC019541 Drosophil
35	17.4	82.9	122156	2	AC106173 Rattus no
36	17.4	82.9	170900	3	AC107487 Drosophil
37	17.4	82.9	171451	3	AC010561 Drosophil
38	17.4	82.9	189613	2	AC118117 Rattus no
39	17.4	82.9	306945	3	AE003476 Drosophil
40	17.4	82.9	309233	2	AC098557 Rattus no
41	17	81.0	119723	2	AC111920 Rattus no
42	17	81.0	173184	9	AC112515 Homo sapi
43	17	81.0	236018	2	AC115296 Mus muscu
44	17	81.0	251938	2	AC124763 Mus muscu
45	16.8	80.0	11937	8	ATBETAFRU X99111 A.thaliana

ALIGNMENTS

RESULT 1	AX014718	AX014718	Sequence 18 from Patent WO9953077	21 bp	DNA	linear	PAT 07-SEP-2000
LOCUS	AX014718	Sequence 18 from Patent WO9953077					
DEFINITION	AX014718	Sequence 18 from Patent WO9953077					
ACCESSION	AX014718	Sequence 18 from Patent WO9953077					
VERSION	AX014718.1	GI:10040991					
KEYWORDS		synthetic construct.					
SOURCE		synthetic construct					
ORGANISM		artificial sequences					
REFERENCE		1 (bases 1 to 21)					
AUTHORS		Schwartz J.C., Gros C., Ouimet T., Rose C., Bonhomme M.C. and					
TITLE		Novel nep ii membrane metalloprotease and its use for screening					
		inhibitors useful in therapy					

Db 1760 CCGAAGTTCCTTGAGGCTCC 1740

RESULT 3  
AC094732/C  
LOCUS  
DEFINITION

AC094732 174953 bp DNA linear HTG 20-DEC-2001  
Rattus norvegicus clone CH230-516, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
56 unordered pieces.

AC094732  
AC094732.2 GI:17941511  
HTG: HTGS\_PRAISEL  
Rattus norvegicus  
Rattus norvegicus  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

1 (bases 1 to 174953)

REFERENCE  
AUTHORS  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,  
Benton,J., Blinige,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
Bowie,S., Brivaga,M., Brown,B., Brown,M., Bryant,N.P., Buhay,C.,  
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,  
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,  
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,  
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,  
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,  
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,  
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,  
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,  
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,  
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S.,  
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,J.B., Korvah,J.,  
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,  
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,  
Loulseghe,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,  
Meig,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,  
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,  
Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,  
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,  
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N.,  
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,  
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,  
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,  
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R.,  
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,  
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.

Direct Submission  
Unpublished  
2 (bases 1 to 174953)  
Worley K.C.  
Direct Submission  
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Dec 20, 2001 this sequence version replaced gi:15624568.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
-----  
Center project name: GBGF

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

```

Center clone name: CH230-516
----- Summary Statistics -----
findPhrapList
Assembly program: Phrap; version 0.990329Ffirst call to
Consensus quality: 152255 bases at least Q40
Consensus quality: 158448 bases at least Q30
Consensus quality: 164461 bases at least Q20
Estimated insert size: 155965; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 2.2x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 56 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1.
13782: contig of 13782 bp in length
13783: 13882: gap of unknown length
13883: 23287: contig of 9405 bp in length
23288: 23387: gap of unknown length
23388: 28081: contig of 4694 bp in length
28082: 28181: gap of unknown length
28182: 33807: contig of 5626 bp in length
33808: 33907: gap of unknown length
33908: 39271: contig of 5364 bp in length
39272: 39371: gap of unknown length
39372: 44271: contig of 4899 bp in length
44271: 44370: gap of unknown length
44371: 47723: contig of 3353 bp in length
47724: 47823: gap of unknown length
47824: 53427: contig of 5604 bp in length
53428: 53527: gap of unknown length
53528: 57303: contig of 3776 bp in length
57304: 57403: gap of unknown length
57404: 62018: contig of 4615 bp in length
62019: 62118: gap of unknown length
62119: 66576: contig of 4558 bp in length
66577: 66776: gap of unknown length
66777: 70201: contig of 3425 bp in length
70202: 70301: gap of unknown length
70302: 74082: contig of 3781 bp in length
74083: 74182: gap of unknown length
74183: 78306: contig of 4124 bp in length
78307: 78406: gap of unknown length
78407: 81222: contig of 2816 bp in length
81223: 81322: gap of unknown length
81323: 84552: contig of 3230 bp in length
84553: 84652: gap of unknown length
84653: 88339: contig of 4187 bp in length
88339: 88939: gap of unknown length
88940: 92256: contig of 3317 bp in length
92257: 92356: gap of unknown length
92357: 94995: contig of 2639 bp in length
94996: 95095: gap of unknown length
95096: 98465: contig of 3370 bp in length
98466: 98565: gap of unknown length
98566: 101632: contig of 3067 bp in length
101633: 101732: gap of unknown length
101733: 104939: contig of 3207 bp in length
104940: 105039: gap of unknown length
105040: 109248: contig of 4209 bp in length
109249: 109348: gap of unknown length
109349: 112435: contig of 3087 bp in length
112436: 112535: gap of unknown length
112536: 114576: contig of 2141 bp in length
114577: 114776: gap of unknown length
114777: 117567: contig of 2791 bp in length
117568: 117667: gap of unknown length

```

```

117568 119209: contig of 1542 bp in length
119210 119309: gap of unknown length
121797 121797: contig of 2488 bp in length
121798 121897: gap of unknown length
121898 124293: contig of 2396 bp in length
124294 124393: gap of unknown length
124394 126428: contig of 2035 bp in length
126429 126528: gap of unknown length
126529 129525: contig of 2997 bp in length
129526 129625: gap of unknown length
129626 132334: contig of 2709 bp in length
132335 132434: gap of unknown length
132435 135274: contig of 2840 bp in length
135275 135374: gap of unknown length
135375 138074: contig of 2700 bp in length
138075 138174: gap of unknown length
138175 139985: contig of 1811 bp in length
139986 140085: gap of unknown length
140086 142273: contig of 2188 bp in length
142274 142373: gap of unknown length
142374 143598: contig of 1225 bp in length
143599 145434: contig of 1736 bp in length
145435 145534: gap of unknown length
145535 146985: contig of 1451 bp in length
146986 147085: gap of unknown length
147086 148099: contig of 1014 bp in length
148100 148199: gap of unknown length
148200 150915: contig of 2716 bp in length
150916 151015: gap of unknown length
151016 152501: contig of 1486 bp in length
152502 152601: gap of unknown length
152602 154010: contig of 1409 bp in length
154011 154110: gap of unknown length
154111 155758: contig of 1648 bp in length
155759 155858: gap of unknown length
155859 157622: contig of 1764 bp in length
157623 157722: gap of unknown length
157723 159428: contig of 1706 bp in length
159429 159528: gap of unknown length
159529 161209: contig of 1681 bp in length
161210 161309: gap of unknown length
161310 163413: contig of 2104 bp in length
163414 163513: gap of unknown length
163514 164702: contig of 1189 bp in length
164703 164802: gap of unknown length
164803 165998: contig of 1196 bp in length
165999 166098: gap of unknown length
166099 167412: contig of 1314 bp in length
167413 167512: gap of unknown length
167513 169231: contig of 1719 bp in length
169232 169331: gap of unknown length
169332 170534: contig of 1203 bp in length
170535 170634: gap of unknown length
170635 172047: contig of 1413 bp in length
172048 172147: gap of unknown length
172148 173509: contig of 1362 bp in length
173510 173609: gap of unknown length
173610 174953: contig of 1344 bp in length.

```

## FEATURES

Location/Qualifiers

Query Match 100.0%; Score 21; DB 2; Length 174953;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGAAGTTTCTTGAGGCTCC 21  
 |||||  
 Db 64550 CCGAAGTTTCTTGAGGCTCC 64530

RESULT 4  
 AF302075/c  
 LOCUS

AF302075

2583 bp

mRNA

linear

ROD 11-JUN-2001



```

DEFINITION Mus musculus neprilysin-like peptidase alpha mRNA, complete cds.
ACCESSION AF302075
KEYWORDS AF302075.1 GI:10505359
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 2583)
AUTHORS Shirotsani,K., Tsubuki,S., Iwata,N., Takaki,Y., Harigaya,W.,
Iwatsubo,T., Kiryu-Seo,S., Kiyama,H., Iwata,H., Tomita,T.,
Iwatsubo,T. and Saido,T.C.
TITLE Neprilysin degrades both amyloid beta peptides 1-40 and 1-42 most
rapidly and efficiently among thiorphan- and
phosphoramidon-sensitive endopeptidases
J. Biol. Chem. 276 (24), 21895-21901 (2001)
JOURNAL 21293028
MEDLINE 11278416
PUBMED
REFERENCE 2 (bases 1 to 2583)
AUTHORS Shirotsani,K. and Saido,T.C.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-2000) Proteolytic Neuroscience Laboratory, Brain
Science Institute, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama
351-0198, Japan
FEATURES
source Location/Qualifiers
1..2583
/organism="Mus musculus"
/db_xref="taxon:10090"
25..2253
/feature="endopeptidase"
/codon_start=1
/product="neprilysin-like peptidase alpha"
/protein_id="AAG18446.1"
/db_xref="GI:10505360"
/translation="MVERAGWCRKSPGFVEYGLMVLVLLLLGAILVTLGVFYSIALRD
SSLSKSDICTTPSCVIAAARILENNQSRNCFYQACGWLHRHVIPTNSRYSVF
DILRDEVLTKGVLEDSTOHRPAVEKAKTLVRSQNMOSVIEKRSEPLLSVLKMGV
GWPVADKWEINMGKLEWELROLAVLNQSFNRVLIDLFWINDQNSRHVLIIDQPT
LGMPSREYFQEDNNHKVRKAYLEPMTSVATMLRQDNLSKESAMVREMAEVELEET
HLANATVPQERKRDVTLVYHRMDLMELQERGLKGFNTLFIQNVLSSVEVLPDDEE
VYVYGPYLENLEDIIDSYSARTQMNYLVRLVLDRLGSLQSFKEARVDYKALYGT
TVVEVWRKCVSVNMSMESAVGSLYIKRAFSDKSTVRELIEKRSFYEDLYFENGLOLNLK
WNEDESKKAQKAMNIREQIGYPDYILEDNNKHLDEEYSSLTFFEDLYFENGLOLNLK
NNAQSLKLRKQDNQWLIIGAAVNVAFYSPNRNQLVFPAGILOPPFFSKDQPSLN
LADNQNVNGFSTLGENIADNGVRQAYKAYLRWLADGDKQDRLPGLNLTLYAQLFIFY
AQVWCGSYRPEFAVQSIKTDVHSPKRYVLGSLQNLPGFSEAFHCPGRGSPMHPMKRCR
IW"
BASE COUNT 665 a 667 c 736 g 515 t
ORIGIN
Query Match 92.4%; Score 19.4; DB 10; Length 2583;
Best Local Similarity 95.2%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCCGAAGTTCTTGAGGCTCC 21
||||| |||||||
Db 1582 CCCGAAGTTCTTGAGGCTCC 1562

RESULT 5
AF157106/c
LOCUS AF157106 2601 bp mRNA linear ROD 25-NOV-1999
DEFINITION Mus musculus soluble secreted endopeptidase delta mRNA,
alternatively spliced product, complete cds.
ACCESSION AF157106
VERSION AF157106.1 GI:6467400
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 2601)
AUTHORS Shirotsani,K. and Saido,T.C.
TITLE Neprilysin degrades both amyloid beta peptides 1-40 and 1-42 most
rapidly and efficiently among thiorphan- and
phosphoramidon-sensitive endopeptidases
J. Biol. Chem. 276 (24), 21895-21901 (2001)
JOURNAL 21293028
MEDLINE 11278416
PUBMED
REFERENCE 2 (bases 1 to 2601)
AUTHORS Shirotsani,K. and Saido,T.C.

```

```

AUTHORS Ikeda,K., Emoto,N., Raharjo,S.B., Nurfantari,Y., Saiki,K.,
Yokoyama,M. and Matsuo,M.
TITLE Molecular identification and characterization of novel
membrane-bound metalloprotease, the soluble secreted form of which
hydrolyzes a variety of vasoactive peptides
J. Biol. Chem. 274 (45), 32469-32477 (1999)
JOURNAL 20011457
MEDLINE 10542292
PUBMED
REFERENCE 2 (bases 1 to 2601)
AUTHORS Ikeda,K., Emoto,N. and Matsuo,M.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-1999) International Center for Medical Research,
Kobe University School of Medicine, 7-5-1 Kusunoki, Chuoh, Kobe
6500017, Japan
FEATURES
source Location/Qualifiers
1..2601
/organism="Mus musculus"
/db_xref="taxon:10090"
59..2287
/feature="SEP(delta); metalloprotease; alternatively spliced"
/codon_start=1
/product="soluble secreted endopeptidase delta"
/protein_id="AAFI3153.1"
/db_xref="GI:6467401"
/translation="MVERAGWCRKSPGFVEYGLMVLVLLLLGAILVTLGVFYSIALRD
SSLSKSDICTTPSCVIAAARILENNQSRNCFYQACGWLHRHVIPTNSRYSVF
DILRDEVLTKGVLEDSTOHRPAVEKAKTLVRSQNMOSVIEKRSEPLLSVLKMGV
GWPVADKWEINMGKLEWELROLAVLNQSFNRVLIDLFWINDQNSRHVLIIDQPT
LGMPSREYFQEDNNHKVRKAYLEPMTSVATMLRQDNLSKESAMVREMAEVELEET
HLANATVPQERKRDVTLVYHRMDLMELQERGLKGFNTLFIQNVLSSVEVLPDDEE
VYVYGPYLENLEDIIDSYSARTQMNYLVRLVLDRLGSLQSFKEARVDYKALYGT
TVVEVWRKCVSVNMSMESAVGSLYIKRAFSDKSTVRELIEKRSFYEDLYFENGLOLNLK
WNEDESKKAQKAMNIREQIGYPDYILEDNNKHLDEEYSSLTFFEDLYFENGLOLNLK
NNAQSLKLRKQDNQWLIIGAAVNVAFYSPNRNQLVFPAGILOPPFFSKDQPSLN
LADNQNVNGFSTLGENIADNGVRQAYKAYLRWLADGDKQDRLPGLNLTLYAQLFIFY
AQVWCGSYRPEFAVQSIKTDVHSPKRYVLGSLQNLPGFSEAFHCPGRGSPMHPMKRCR
IW"
BASE COUNT 655 a 681 c 748 g 517 t
ORIGIN
Query Match 92.4%; Score 19.4; DB 10; Length 2601;
Best Local Similarity 95.2%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCCGAAGTTCTTGAGGCTCC 21
||||| |||||||
Db 1616 CCCGAAGTTCTTGAGGCTCC 1596

RESULT 6
AF302076/c
LOCUS AF302076 2652 bp mRNA linear ROD 11-JUN-2001
DEFINITION Mus musculus neprilysin-like peptidase beta mRNA, complete cds.
ACCESSION AF302076
VERSION AF302076.1 GI:10505361
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 2652)
AUTHORS Shirotsani,K., Tsubuki,S., Iwata,N., Takaki,Y., Harigaya,W.,
Matsuo,M., Kiryu-Seo,S., Kiyama,H., Iwata,H., Tomita,T.,
Iwatsubo,T. and Saido,T.C.
TITLE Neprilysin degrades both amyloid beta peptides 1-40 and 1-42 most
rapidly and efficiently among thiorphan- and
phosphoramidon-sensitive endopeptidases
J. Biol. Chem. 276 (24), 21895-21901 (2001)
JOURNAL 21293028
MEDLINE 11278416
PUBMED
REFERENCE 2 (bases 1 to 2652)
AUTHORS Shirotsani,K. and Saido,T.C.

```



RIGLSQRFKARVDYRKALYGTVEVRRECYSVNSNMESAVGSLYIKRAFSSKDS  
KSTVRELIEKRSVFNDELNMDESKKAKOAKMNIHQIGYIPDYILEDNNKHL  
DEEYSSTFTFEDLYFENGQNLKNAQSLKLEKVDONLWIIIGAAVNAFYSPNRN  
OIVFPAGILQPPFSKQPOQSLNFGGIGWVIGHEITHGFDNGRNFKNMGLDWSN  
FSARHFOQSQCMIIYOYGNFSEWELADNQNVNGFSLGENIADNGVRQAYKAYLRWLA  
DGGKQDRLPLGLNTYAQLFFINYAQVWCGSYRPEFAVQSIKTDVHSPKRYRVLGSLQ  
LPGFSEAFHCPRGSPHMKRCRIW"

BASE COUNT 717 a 774 c 827 g 574 t

ORIGIN

Query Match 92.4%; Score 19.4; DB 10; Length 2892;

Best Local Similarity 95.2%; Pred. No. 11;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCGAAGTTCTTGAGGCTCC 21

||||| ||||| ||||| ||||| |||||

Db 1907 CCCGAAGTTCTTGAGGCTCC 1887

RESULT 9

AX033272/c

LOCUS

DEFINITION Sequence 12 from Patent WO0047750.

ACCESSION AX033272

VERSION AX033272.1 GI:10280087

KEYWORDS

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE 1 (bases 1 to 2925)

JOURNAL Boileau, G. and Desgroselliers, L.

Patent: WO 0047750-A 12 17-AUG-2000;

BOILEAU GUY (CA); DESGROSSELLERS LUC (CA); UNIVERSITE DE MONTREAL

(CA)

FEATURES

source Location/Qualifiers

1..2925

/organism="Mus musculus"

/db\_xref="taxon:10090"

332..2629

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC09977.1"

/db\_xref="GI:10280088"

/translation="MVERAGWCKKSPGFVEYGLMVLILLGAIIVTLGVFYSIGKQL

PLLTSLHFSWDETVVVKRALRDSLSKSDICTTPSCVIAAARILENMDQSRNCPENFY

QYACGGLRHVHPETNRSYVFDILRDEVLKGVLESDTSQHRPAVEKATLYRS

CMQSVIERKDSRSEPLSLVKMVGWVPMVKNETMGLKWELEQLAVLNSQFNRRVL

IDLFWDDQSSRHVYIDQPTLGMPSREYFQEDNNHKVRKAYLEFMTSVATMLRK

DONLSKESAMVREMAEVELETHLANATVPOKRHDVTALYHRMDLMEQLERFLKG

FNWTLFIQNVLSSEVELFPDEEVVYGYPIYLENLEDIIDSYSARTMQLYVRLVD

RIGLSQRFKARVDYRKALYGTVEVRRECYSVNSNMESAVGSLYIKRAFSSKDS

KSTVRELIEKRSVFNDELNMDESKKAKOAKMNIHQIGYIPDYILEDNNKHL

DEEYSSTFTFEDLYFENGQNLKNAQSLKLEKVDONLWIIIGAAVNAFYSPNRN

QIVFPAGILQPPFSKQPOQSLNFGGIGWVIGHEITHGFDNGRNFKNMGLDWSN

FSARHFOQSQCMIIYOYGNFSEWELADNQNVNGFSLGENIADNGVRQAYKAYLRWLA

DGGKQDRLPLGLNTYAQLFFINYAQVWCGSYRPEFAVQSIKTDVHSPKRYRVLGSLQ

LPGFSEAFHCPRGSPHMKRCRIW"

BASE COUNT 710 a 797 c 836 g 582 t

ORIGIN

Query Match 92.4%; Score 19.4; DB 6; Length 2925;

Best Local Similarity 95.2%; Pred. No. 11;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCGAAGTTCTTGAGGCTCC 21

||||| ||||| ||||| ||||| |||||

Db 1958 CCCGAAGTTCTTGAGGCTCC 1938

RESULT 10

AF176569/c

LOCUS

DEFINITION

ACCESSION AF176569

VERSION AF176569.1 GI:7769082

KEYWORDS

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE 1 (bases 1 to 2925)

JOURNAL Ghaddar, G., Ruchon, A. F., Carpentier, M., Marcinkiewicz, M.,

Seidah, N. G., Crine, P., Desgroselliers, L. and Boileau, G.

Molecular cloning and biochemical characterization of a new mouse

testis soluble-zinc-metalloproteinase of the neprilysin family

Biochem. J. 347 (Pt 2), 419-429 (2000)

MEDLINE 21060448

PUBMED 10749671

REFERENCE 2 (bases 1 to 2925)

AUTHORS Ghaddar, G., Ruchon, A. F., Desgroselliers, L. and Boileau, G.

Submitted (09-AUG-1999) Biochemistry, University of Montreal, 2900

Edouard Montpetit, Montreal, Que H3T 1J4, Canada

Location/Qualifiers

1..2925

/organism="Mus musculus"

/db\_xref="taxon:10090"

/tissue\_type="testis"

1..2925

/gene="Nll"

332..2629

/note="Nll: metallo-endopeptidase"

/codon\_start=1

/product="neprilysin-like metalloproteinase 1"

/protein\_id="AAF69247.1"

/db\_xref="GI:7769083"

/translation="MVERAGWCKKSPGFVEYGLMVLILLGAIIVTLGVFYSIGKQL

PLLTSLHFSWDETVVVKRALRDSLSKSDICTTPSCVIAAARILENMDQSRNCPENFY

QYACGGLRHVHPETNRSYVFDILRDEVLKGVLESDTSQHRPAVEKATLYRS

CMQSVIERKDSRSEPLSLVKMVGWVPMVKNETMGLKWELEQLAVLNSQFNRRVL

IDLFWDDQSSRHVYIDQPTLGMPSREYFQEDNNHKVRKAYLEFMTSVATMLRK

DONLSKESAMVREMAEVELETHLANATVPOKRHDVTALYHRMDLMEQLERFLKG

FNWTLFIQNVLSSEVELFPDEEVVYGYPIYLENLEDIIDSYSARTMQLYVRLVD

RIGLSQRFKARVDYRKALYGTVEVRRECYSVNSNMESAVGSLYIKRAFSSKDS

KSTVRELIEKRSVFNDELNMDESKKAKOAKMNIHQIGYIPDYILEDNNKHL

DEEYSSTFTFEDLYFENGQNLKNAQSLKLEKVDONLWIIIGAAVNAFYSPNRN

QIVFPAGILQPPFSKQPOQSLNFGGIGWVIGHEITHGFDNGRNFKNMGLDWSN

FSARHFOQSQCMIIYOYGNFSEWELADNQNVNGFSLGENIADNGVRQAYKAYLRWLA

DGGKQDRLPLGLNTYAQLFFINYAQVWCGSYRPEFAVQSIKTDVHSPKRYRVLGSLQ

LPGFSEAFHCPRGSPHMKRCRIW"

BASE COUNT 710 a 797 c 836 g 582 t

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

Qy

Db

RESULT 11

AC120432/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mus musculus clone RP24-553K14, LOW-PASS SEQUENCE SAMPLING.

AC120432

AC120432.1

GI:20455766

HTG; HTGS\_PHASE0.

Mus musculus

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AC120432

Mus musculus clone RP24-553K14, LOW-PASS SEQUENCE SAMPLING.

AC120432

GI:20455766

HTG; HTGS\_PHASE0.

Mus musculus

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 67142)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Mus musculus, clone RP24-553K14  
Unpublished

2 (bases 1 to 67142)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,  
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,  
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faros,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kellis,C., LaRocque,K., Lamazares,R.,  
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,  
Minova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
COMMENT

Submitted (06-MAY-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www.seq.wi.mit.edu>  
Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L26051  
Center clone name: 553\_K\_14

NOTE: This record contains 82 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1  
692 791: contig of 691 bp in length  
792 1524: contig of 733 bp in length  
1525 1624: gap of 100 bp  
1625 2355: contig of 731 bp in length  
2356 2455: gap of 100 bp  
2456 3135: contig of 680 bp in length  
3136 3235: gap of 100 bp  
3236 3928: contig of 693 bp in length  
3929 4028: gap of 100 bp  
4029 4730: contig of 702 bp in length  
4731 4830: gap of 100 bp  
4831 5532: contig of 702 bp in length  
5533 5632: gap of 100 bp  
5633 6372: contig of 740 bp in length  
6373 6472: gap of 100 bp  
6473 7212: contig of 740 bp in length  
7213 7312: gap of 100 bp  
7313 8009: contig of 697 bp in length

8010 8109: gap of 100 bp  
8110 8831: contig of 722 bp in length  
8832 8931: gap of 100 bp  
8932 9631: contig of 700 bp in length  
9632 9731: gap of 100 bp  
9732 10448: contig of 717 bp in length  
10449 10548: gap of 100 bp  
10549 11268: contig of 720 bp in length  
11269 11368: gap of 100 bp  
11369 12087: contig of 719 bp in length  
12088 12187: gap of 100 bp  
12188 12931: contig of 744 bp in length  
12932 13031: gap of 100 bp  
13032 13780: contig of 749 bp in length  
13781 13880: gap of 100 bp  
13881 14576: contig of 696 bp in length  
14577 14676: gap of 100 bp  
14677 15410: contig of 734 bp in length  
15411 15510: gap of 100 bp  
15511 16248: contig of 738 bp in length  
16249 16348: gap of 100 bp  
16349 17053: contig of 705 bp in length  
17054 17153: gap of 100 bp  
17154 17836: contig of 683 bp in length  
17837 17936: gap of 100 bp  
17937 18651: contig of 715 bp in length  
18652 18751: gap of 100 bp  
18752 19477: contig of 726 bp in length  
19478 19577: gap of 100 bp  
19578 20282: contig of 705 bp in length  
20283 20382: gap of 100 bp  
20383 21125: contig of 743 bp in length  
21126 21225: gap of 100 bp  
21226 21974: contig of 749 bp in length  
21975 22074: gap of 100 bp  
22075 22783: contig of 709 bp in length  
22784 22883: gap of 100 bp  
22884 23584: contig of 701 bp in length  
23585 23684: gap of 100 bp  
23685 24409: contig of 725 bp in length  
24410 24509: gap of 100 bp  
24510 25220: contig of 711 bp in length  
25221 25320: gap of 100 bp  
25321 26029: contig of 709 bp in length  
26030 26129: gap of 100 bp  
26130 26849: contig of 720 bp in length  
26850 26949: gap of 100 bp  
26950 27674: contig of 725 bp in length  
27675 27774: gap of 100 bp  
27775 28509: contig of 735 bp in length  
28510 28609: gap of 100 bp  
28610 29341: contig of 732 bp in length  
29342 29441: gap of 100 bp  
29442 30176: contig of 735 bp in length  
30177 30276: gap of 100 bp  
30277 31015: contig of 739 bp in length  
31016 31115: gap of 100 bp  
31116 31844: contig of 729 bp in length  
31845 31944: gap of 100 bp  
31945 32639: contig of 695 bp in length  
32640 32739: gap of 100 bp  
32740 33484: contig of 745 bp in length  
33485 33584: gap of 100 bp  
33585 34322: contig of 738 bp in length  
34323 34422: gap of 100 bp  
34433 35120: contig of 698 bp in length  
35121 35220: gap of 100 bp  
35221 35918: contig of 698 bp in length  
35919 36018: gap of 100 bp  
36019 36720: contig of 702 bp in length  
36721 36820: gap of 100 bp  
36821 37533: contig of 719 bp in length  
37540 37639: gap of 100 bp

```

* 37640 38360: contig of 721 bp in length
* 38361 38460: gap of 100 bp
* 38461 39198: contig of 738 bp in length
* 39199 39298: gap of 100 bp
* 39299 40038: contig of 740 bp in length
* 40039 40138: gap of 100 bp
* 40139 40882: contig of 744 bp in length
* 40883 40982: gap of 100 bp
* 40983 41707: contig of 725 bp in length
* 41708 41807: gap of 100 bp
* 41808 42513: contig of 706 bp in length
* 42514 42613: gap of 100 bp
* 42614 43325: contig of 712 bp in length
* 43326 43425: gap of 100 bp
* 43426 44133: contig of 708 bp in length
* 44134 44233: gap of 100 bp
* 44234 44954: contig of 721 bp in length
* 44955 45054: gap of 100 bp
* 45055 45786: contig of 732 bp in length
* 45787 45886: gap of 100 bp
* 45887 46618: contig of 732 bp in length
* 46619 46718: gap of 100 bp
* 46719 47455: contig of 737 bp in length
* 47456 47555: gap of 100 bp
* 47556 48296: contig of 741 bp in length
* 48297 48396: gap of 100 bp
* 48397 49125: contig of 729 bp in length
* 49126 49225: gap of 100 bp
* 49226 49974: contig of 749 bp in length
* 49975 50074: gap of 100 bp
* 50075 50768: contig of 694 bp in length
* 50769 50868: gap of 100 bp
* 50869 51593: contig of 725 bp in length
* 51594 51693: gap of 100 bp
* 51694 52421: contig of 728 bp in length
* 52422 52521: gap of 100 bp
* 52522 53241: contig of 720 bp in length
* 53242 53341: gap of 100 bp
* 53342 54030: contig of 689 bp in length
* 54031 54130: gap of 100 bp
* 54131 54832: contig of 702 bp in length
* 54833 54932: gap of 100 bp
* 54933 55635: contig of 703 bp in length
* 55636 55735: gap of 100 bp
* 55736 56426: contig of 691 bp in length
* 56427 56526: gap of 100 bp

```

Query Match 92.4%; Score 19.4; DB 2; Length 67142;

Best Local Similarity 95.2%; Pred. No. 13;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 CCCGAAGTTCTTGAGGCTCC 21
    ||||| ||||| ||||| |||||
DB 22390 CCCGAAGTTCTTGAGGCTCC 22370

```

# RESULT 12

AL607032/c

LOCUS

DEFINITION AL607032 208249 bp DNA linear HTG 24-JUL-2002

PROGRESS \*\*\*, 2 unordered pieces.

ACCESSION AL607032

VERSION AL607032.15 GI:21955491

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_ACTIVIFIN; HTGS\_DRAFT; HTGS\_FULLTOP.

SOURCE Mus musculus.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 208249)

Tracey, A.

Direct Submission

TITLE Submitted (23-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,

JOURNAL Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

## COMMENT

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Jul 25, 2002 this sequence version replaced gi:21627906.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: BM15L19

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator; 11% of reads

Chemistry: Dye-terminator Big Dye; 88% of reads

Consensus quality: 207834 bases at least Q40

Consensus quality: 207956 bases at least Q30

Consensus quality: 208060 bases at least Q20

Insert size: 219051; 2.3% error; agarose-fp

Quality coverage: 12.41x in Q20 bases; sum-of-contigs Quality

coverage: 12.08x in Q20 bases; agarose-fp

-----

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 2 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 137303: contig of 137303 bp in length

\* 137304 137403: gap of 100 bp

\* 137404 208249: contig of 70846 bp in length.

## FEATURES

source

1..208249

Location/Qualifiers

/organism="Mus musculus"

/db\_xref="taxon:10090"

/chromosome="4"

/clone="RP23-15L19"

/clone\_lib="RPCI-23"

1..137303

/note="assembly\_fragment:04071"

/note="assembly\_fragment:01838"

clone\_end:77

vector\_side:right"

BASE COUNT 54098 a 53647 c 51328 g 49076 t 100 others

## ORIGIN

```

Query Match 92.4%; Score 19.4; DB 2; Length 208249;
Best Local Similarity 95.2%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 CCCGAAGTTCTTGAGGCTCC 21
    ||||| ||||| ||||| |||||
DB 180533 CCCGAAGTTCTTGAGGCTCC 180513

```

## RESULT 13

TBCHRIA\_06

WPCOMMENT

Sequence split into 10 fragments LOCUS TBCHRIA Accession AL359782

Fragment Name Begin End

TBCHRIA\_00 1 110000

TBCHRIA\_01 100001 210000

TBCHRIA\_02 200001 310000

TBCHRIA\_03 300001 410000

TBCHRIA\_04 400001 510000

TBCHRIA\_05 500001 610000

TBCHRIA\_06 600001 710000

TBCHRIA\_07 700001 810000

TBCHRIA\_08 800001 910000

TBCHRIA\_09 900001 941631

Continuation (7 of 10) of TBCHRLA from base 600001 (AL359782 Trypanosoma brucei chromosome 1)

Query Match 87.6%; Score 18.4; DB 2; Length 110000;  
Best Local Similarity 95.0%; Pred. No. 49;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCGAAGTTTCTTGAGGCTC 20  
|||||

Db 75627 CCCGAAGTTTCTTGAGGCTC 75646  
|||||

## RESULT 14

AX146976/c  
LOCUS AX146976 2076 bp DNA linear PAT 08-JUN-2001  
DEFINITION Sequence 1 from Patent WO0136610.  
ACCESSION AX146976  
VERSION AX146976.1 GI:14346247  
KEYWORDS  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 2076)  
Deleersnijder, W., Wiegers R. and Weske, M.

TITLE Human enzymes of the metalloprotease family

JOURNAL Patent: WO 0136610-A 1 25-MAY-2001;

Solvay Pharmaceuticals B.V. (NL)

Location/Qualifiers

1..2076

/organism="Homo sapiens"

/db\_xref="taxon:9606"

<1..2076

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC41159.1"

/db\_xref="GI:14346248"

/translation="CTTPGCVIAAARILQNMPTTECDDFYQFACGGWLRHHVIPET  
NSRISFDVLEVLKAVLENSTAKDRPAVEKARTLYRSCMNQSVIEKRGSPQLL  
DILEVGGWPMVMDRWNETVGLWELEQLALMNSQFNRRVLIDLFVNDONSRRHI  
IYIDPTLGMPSREYFNGGSRKRYEAYLQFVSVATILLREDANLFRDSCLVQEDMM  
QVLELTQLAKATVPQERHDVIALYHMGLEEQSQFGLKGFNWLFIQTIVLSSVKI  
KLPLDEVVYIGIPILQNLNI IDTSARTIQNYLVWRLVLDRIKSLRSQRFKPTRVNY  
KRALFTGMEVREVCVGVNSNMENAVGSLYVREAFPGDSDKSMVRELIDKVRTVY  
ETLDELGMDEESKKAQEKMSIREQIGHDPDYILEMNRRLDEEYNLNFSEDLYPE  
NSLQNLKVGQAQSLRLKREKVDNLTIGAAVNVAFYSPNRNOIVFPAQILQPPFFSK  
EPCALNFGGIGVIGHEITHGFDNCRFDKNGMDWNSFSTQFRQSECMYIO  
YGNYSNDLADEQVNGFNILGENIADNGGVQAYKAYLKWAEKGDQQLPGLDLTHE  
QLFFINVAQVWCGSYRPEFAIQSIKTDVHSPLRYRLVLSLQNLAAAFADTFHCARGTPM  
HPKRCRVW"

BASE COUNT 493 a 578 c 636 g 369 t

## ORIGIN

Query Match 84.8%; Score 17.8; DB 6; Length 2076;

Best Local Similarity 90.5%; Pred. No. 89;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCGAAGTTTCTTGAGGCTCC 21  
|||||

Db 1405 CCCGAAGTTTCTTGAGGCTCC 1385  
|||||

## RESULT 15

AX319864/c  
LOCUS AX319864 2232 bp DNA linear PAT 14-DEC-2001  
DEFINITION Sequence 28 from Patent WO0183782.  
ACCESSION AX319864  
VERSION AX319864.1 GI:17901454  
KEYWORDS  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1

AUTHORS Ploymann, G.D., Whyte, D., Sudarsanam, S., Manning, G., Caenepeel, S. and Payne, V.

TITLE Novel proteases

JOURNAL Patent: WO 0183782-A 28 08-NOV-2001;

FEATURES Sugem, Inc. (US)

Location/Qualifiers

1..2232

/organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT 512 a 620 c 705 g 395 t

## ORIGIN

Query Match 84.8%; Score 17.8; DB 6; Length 2232;

Best Local Similarity 90.5%; Pred. No. 89;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCGAAGTTTCTTGAGGCTCC 21  
|||||

Db 1561 CCCGAAGTTTCTTGAGGCTCC 1541  
|||||

Search completed: July 8, 2003, 03:35:06

Job time : 234.102 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: July 8, 2003, 00:43:28 ; Search time 131.941 Seconds  
(without alignments)  
358.431 Million cell updates/sec

Title: US-09-647-780A-18  
Perfect score: 21  
Sequence: 1 cccgaagtttcttgaagctcc 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs; 1125999159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	21	AAZ28824	Rat membrane metal
2	21	100.0	20	AAZ28810	Rat membrane metal
3	19.4	92.4	2286	AD28130	Soluble secreted e
4	19.4	92.4	2925	AAA63763	CDNA encoding neut
5	17.8	84.8	1746	ABV23433	Human prostate exp
6	17.8	84.8	1746	ABV29290	Human prostate exp
7	17.8	84.8	2076	AAF89737	Nucleotide sequenc
8	17.8	84.8	2232	AAZ97186	Human metalloprote
9	17.8	84.8	2262	AAF89739	Nucleotide sequenc

C 10	17.8	84.8	2318	24	AAZ30580	Human protease, PR
C 11	17.8	84.8	2340	22	AAZ89738	Nucleotide sequenc
C 12	17.8	84.8	2580	24	ABN84280	Human SEP endopept
C 13	17.8	84.8	2636	22	AAF59660	Human neprilysin-1
C 14	17.8	84.8	2663	22	AAF59661	Human neprilysin-1
C 15	17.8	84.8	2676	21	AAA63764	CDNA encoding neut
C 16	17.8	84.8	2714	22	AAF59659	Human neprilysin-1
C 17	17.8	84.8	2893	24	ABN84379	Human SEP endopept
C 18	17.8	84.8	2893	24	AAZ28544	Human soluble secr
C 19	17.8	84.8	2953	24	ABK48251	CDNA encoding nove
C 20	17.8	84.8	2975	24	AAZ28547	Human SEP cDNA inc
C 21	16.8	80.0	611	21	AAF07838	Fusarium venenatum
C 22	16.8	80.0	2217	23	ABL04527	Drosophila melanog
C 23	16.8	80.0	5728	23	ABL04526	Drosophila melanog
C 24	16.2	77.1	199	19	AAV27369	Streptococcus pneu
C 25	16.2	77.1	199	24	ABQ84837	S. pneumoniae SP05
C 26	16.2	77.1	255	24	ABQ90846	M. capsulatus gene
C 27	16.2	77.1	845	22	AAH06460	Human cDNA clone (
C 28	16.2	77.1	3072	21	AAZ75580	DNA encoding a mou
C 29	16.2	77.1	3072	24	AAZ20688	DNA encoding mouse
C 30	16.2	77.1	3072	24	AAZ22952	Murine soluble zal
C 31	16.2	77.1	3488	19	AAZ96307	S. pneumoniae deri
C 32	16.2	77.1	6204	19	AAV44592	Streptococcus pneu
C 33	16.2	77.1	15213	19	AAV52159	Streptococcus pneu
C 34	16.2	77.1	227968	24	ABK83497	Human cDNA differe
C 35	15.8	75.2	525	22	AAH11538	Human cDNA clone (
C 36	15.8	75.2	992	21	AAZ42424	Arabidopsis thalia
C 37	15.8	75.2	3635	23	ABL04042	Drosophila melanog
C 38	15.8	75.2	4991	23	AAZ75353	DNA encoding novel
C 39	15.8	75.2	20448	23	ABL19989	Drosophila melanog
C 40	15.8	75.2	24971	23	ABL19988	Drosophila melanog
C 41	15.8	75.2	368004	24	ABL57909	Human transporter
C 42	15.4	73.3	353	22	AAH12734	Human cDNA clone (
C 43	15.4	73.3	640	23	ABV02410	Human prostate exp
C 44	15.4	73.3	7240	22	AAH50570	Insulin receptor g
C 45	15.4	73.3	7240	22	AAH31170	Human insulin rece

ALIGNMENTS

RESULT 1  
AAZ28824  
ID AAZ28824 standard; DNA; 21 BP.  
XX  
AC AAZ28824;  
XX  
DT 01-FEB-2000 (first entry)

DE Rat membrane metalloprotease NEPII gene probe. #14.  
KW Rat; membrane metalloprotease; neprilysin II; NEPII; inactivation; ss;  
KW neuron; hormone; peptide messenger; inhibitor; detection; disorder;  
KW cardiovascular disease; neurodegenerative disease; growth disorder;  
KW hypothalamic-hypophyseal axis; endocrine disorder; probe; hybridisation.  
XX Synthetic.  
OS Rattus rattus.  
XX  
XX FR2777291-A1.  
XX  
PD 15-OCT-1999.  
XX  
XX 08-APR-1998; 98FR-0004389.  
XX  
PR 08-APR-1998; 98FR-0004389.  
XX  
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
XX  
PI Ouimet T, Gros C, Haret C, Bonhomme MC, Facchinetti P;  
PI Schwartz JC;  
XX  
DR WPI; 1999-593429/51.



XX New membrane metalloprotease NEP II, involved in proteolysis of  
PT neuronal and hormonal peptides, used to screen for inhibitors;  
PT potentially useful for treating e.g. cardiovascular disease  
XX  
PS Claim 3; Page 23; 29pp; French.  
XX  
CC Sequences AA228811-228827 represent probes for detecting the rat  
CC membrane metalloprotease designated neprilysine II (NEPII) gene  
CC (AA228810). NEPII is involved in (in)activation of neuronal and hormonal  
CC peptide messengers. NEPII is used to screen for specific substrates (used  
CC to detect NEPII in cells and tissues) or inhibitors, which can also be  
CC used to detect NEPII or for treatment of disorders related to peptidergic  
CC signalling in which NEPII is involved, e.g. cardiovascular or  
CC neurodegenerative diseases; growth disorders of endocrine origin;  
CC disturbances of the hypothalamic- hypophyseal axis or endocrine  
CC disorders.  
XX  
XX Sequence 21 BP; 3 A; 7 C; 5 G; 6 T; 0 other;  
SQ  
Query Match 100.0%; Score 21; DB 20; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.29;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCCGAAGTTCTTGAGGCTCC 21  
DB 1 CCCGAAGTTCTTGAGGCTCC 21  
RESULT 2  
AAZ28810/c  
ID AA228810 standard; cDNA; 2765 BP.  
XX  
XX AA228810;  
XX  
XX 01-FEB-2000 (first entry)  
XX  
DE Rat membrane metalloprotease NEPII gene.  
XX  
KW Rat; membrane metalloprotease; neprilysine II; NEPII; inactivation;  
KW neuron; hormone; peptide messenger; inhibitor; detection; disorder;  
KW cardiovascular disease; neurodegenerative disease; growth disorder;  
KW hypothalamic-hypophyseal axis; endocrine disorder; ds.  
XX  
OS Rattus rattus.  
XX  
XX FR2777291-A1.  
XX  
XX 15-OCT-1999.  
XX  
XX 08-APR-1998; 98FR-0004389.  
XX  
XX 08-APR-1998; 98FR-0004389.  
XX  
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
XX  
XX Oulmet T, Gros C, Haret C, Bonhomme MC, Facchinetti P;  
PI Schwartz JC;  
XX  
XX WPI; 1999-593429/51.  
XX  
XX P-PSDB; AAY44177.  
XX  
XX New membrane metalloprotease NEP II, involved in proteolysis of  
PT neuronal and hormonal peptides, used to screen for inhibitors;  
PT potentially useful for treating e.g. cardiovascular disease  
XX  
XX Claim 2; Page 12-16; 29pp; French.  
XX  
PS This sequence represents the gene for the rat membrane metalloprotease  
CC designated neprilysine II (NEPII), which is involved in (in)activation  
CC of neuronal and hormonal peptide messengers. NEPII is used to screen  
CC for specific substrates (used to detect NEPII in cells and tissues) or  
CC inhibitors, which can also be used to detect NEPII or for treatment of

CC disorders related to peptidergic signalling in which NEPII is involved,  
CC e.g. cardiovascular or neurodegenerative diseases; growth disorders of  
CC endocrine origin; disturbances of the hypothalamic-hypophyseal axis or  
CC endocrine disorders.  
XX  
SQ Sequence 2765 BP; 684 A; 735 C; 787 G; 559 T; 0 other;  
Query Match 100.0%; Score 21; DB 20; Length 2765;  
Best Local Similarity 100.0%; Pred. No. 0.55;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCCGAAGTTCTTGAGGCTCC 21  
DB 1760 CCCGAAGTTCTTGAGGCTCC 1740  
RESULT 3  
AAD28130/c  
ID AAD28130 standard; DNA; 2286 BP.  
XX  
XX AAD28130;  
XX  
XX 07-MAY-2002 (first entry)  
XX  
XX Soluble secreted endopeptidase (SEP) consensus DNA.  
XX  
KW Human; secreted endopeptidase; SEP; endocrine; vasotropic; analgesic;  
KW enzyme; gynaecological; antisense-therapy; male erectile dysfunction;  
KW MED; female sexual dysfunction; FSD; female sexual arousal disorder;  
KW FSD; premature ejaculation; anorgasmia; vaginismus; mouse; rat; ds.  
XX  
OS Homo sapiens.  
OS Mus sp.  
OS Rattus sp.  
XX  
XX Key Location/Qualifiers  
FT misc\_feature 1664..2286  
FT /\*tag= a  
FT /note= "Encodes catalytic domain"  
FT  
XX WO200206492-A1.  
XX  
XX 24-JAN-2002.  
XX  
XX 16-JUL-2001; 2001WO-1B01263.  
XX  
XX 14-JUL-2000; 2000GB-0017387.  
XX  
XX (PFIZ ) PFIZER LTD.  
XX (PFIZ ) PFIZER INC.  
XX  
XX Harrow ID, Stacey P, Walsh RT, Wayman CP, Wayman CP, Phillips SC;  
PI WPI; 2002-155042/20.  
XX  
XX An isolated and/or purified nucleic acid encoding a human soluble  
PT secreted endopeptidase which is useful for treating sexual dysfunction,  
PT for e.g. male erectile dysfunction or female sexual dysfunction such as  
PT female sexual arousal disorder -  
XX  
XX Disclosure; Fig 6; 167pp; English.  
XX  
XX The invention relates to an isolated and/or purified nucleic acid  
CC encoding a human soluble secreted endopeptidase (SEP). The anti-SEP  
CC antibody and the compound which inhibits or selectively inhibits the  
CC human SEP protein are useful in the manufacture of a medicament for  
CC the prophylaxis and/or treatment of sexual dysfunction, in particular  
CC male erectile dysfunction (MED) or female sexual dysfunction (FSD)  
CC (e.g. female sexual arousal disorder (FSAD)). The SEP polypeptides are  
CC also useful for treating the above disorders and other disorders such  
CC as premature ejaculation, anorgasmia, sexual pain (e.g. vaginismus) and  
CC hypoaactive sexual desire disorder. The present sequence is SEP consensus  
CC DNA sequence found in human, mouse and rat.



Query Match 92.4%; Score 19.4; DB 21; Length 2925;  
Best Local Similarity 95.2%; Pred. No. 3.7;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCGGAAGTTCTTGAGGCTCC 21  
||||| |||||

DB 1958 CCGGAAGTTCTTGAGGCTCC 1938  
||||| |||||

RESULT 5  
ID ABV23433 standard; CDNA; 1746 BP.  
XX AC ABV23433;  
XX DT  
XX 16-SEP-2002 (first entry)  
XX Human prostate expression marker cDNA 23424.  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
KW OS Homo sapiens.  
XX PN WO200160860-A2.  
XX PD 23-AUG-2001.  
XX PF 20-FEB-2001; 2001WO-US05171.  
XX PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX PI Schlegel R, Endege WO, Monahan JE;  
XX WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
XX Claim 1; Page 4269; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX SQ Sequence 1746 BP; 299 A; 531 C; 580 G; 336 T; 0 other;

Query Match 84.8%; Score 17.8; DB 23; Length 1746;  
Best Local Similarity 90.5%; Pred. No. 23;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGGAAGTTCTTGAGGCTCC 21  
||||| |||||

Db 1525 CCGAAGCTTCTGAGGCTCC 1545

## RESULT 6

ABV29290

ID ABV29290 standard; cDNA; 1746 BP.

XX AC ABV29290;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 29281.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US05171.

XX PR 17-FEB-2000; 2000US-183319P.

XX PR 16-MAR-2000; 2000US-189862P.

XX PR 25-MAY-2000; 2000US-207454P.

XX PR 09-JUN-2000; 2000US-211314P.

XX PR 18-JUL-2000; 2000US-219007P.

XX PR 13-DEC-2000; 2000US-255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of

XX PT prostate cells and correlating with presence of prostate cancer, useful

XX PS Claim 1; Page 6257; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising

XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

XX CC specification or its complement. (I) is useful for:

XX CC (a) assessing whether a patient is afflicted with prostate cancer;

XX CC (b) monitoring the progression of prostate cancer in a patient;

XX CC (c) assessing the efficacy of a test compound to inhibit prostate

XX CC cancer in a patient;

XX CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer

XX CC in a patient;

XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;

XX CC (f) assessing the prostate cell carcinogenic potential of a compound;

XX CC (g) determining whether prostate cancer has metastasized in a patient;

XX CC (h) assessing the aggressiveness or indolence of prostate cancer in a

XX CC patient;

XX CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX SQ Sequence. 1746 BP; 299 A; 531 C; 580 G; 336 T; 0 other;

Query Match 84.8%; Score 17.8; DB 23; Length 1746;

Best Local Similarity 90.5%; Pred. No. 23;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCGAAGTTTCTGAGGCTCC 21

Db 1525 CCGAAGCTTCTGAGGCTCC 1545

## RESULT 7

AAF89737/c

ID AAF89737 standard; DNA; 2076 BP.

XX AC

XX DT 23-JUL-2001 (first entry)

XX DE Nucleotide sequence of a human metalloprotease enzyme IGS5.

XX KW Metalloprotease; IGS5; infection; pain; cancer; diabetes; obesity;

XX KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;

XX KW hypotension; hypertension; urinary retention; osteoporosis;

XX KW angina pectoris; myocardial infarction; stroke; ulcer; allergy;

XX KW benign prostatic hypertrophy; migraine; psychotic disorder;

XX KW neurological disorder; autism; multiple sclerosis; Alzheimer's disease;

XX KW neurodegenerative disease; sleep disorder; epilepsy; kidney disease;

XX KW cardiovascular disease; arteriosclerosis; cerebrovasospasm;

XX KW subarachnoid hemorrhage; cerebral ischemia; cerebral infarction;

XX KW peripheral vascular disease; Raynaud's disease; motility disorder;

XX KW gastrointestinal disorder; gastric emptying; gastroparesis; diarrhoea;

XX KW inflammation; chemotherapy induced injury; tumour invasion;

XX KW immune disorder; arthritis; endotoxin shock; sepsis; diabetes mellitus;

XX KW severe mental retardation; dyskinesia; Huntington's disease;

XX KW Gilles de la Tourette's syndrome; ss.

XX OS Homo sapiens.

XX PN WO200136610-A1.

XX PD 25-MAY-2001.

XX PF 17-NOV-2000; 2000WO-EP11532.

XX PR 19-NOV-1999; 99EP-0203862.

XX PR 31-MAY-2000; 99NL-1013616.

XX PR 31-MAY-2000; 2000EP-0201937.

XX PR 31-MAY-2000; 2000NL-1013356.

XX PA (SOLV ) SOLVAY PHARM BV.

XX PI Deleersnijder W, Wieggers R, Weske M;

XX DR WPI; 2001-343815/36.

XX DR P-PSDB; AAB83840.

XX PT New IGS5 polypeptides useful for treating infections, pain, cancer,

XX PT diabetes, obesity, anorexia, bulimia, asthma, schizophrenia,

XX PT hypertension, urinary retention and Parkinson's disease

XX PS Claim 11; Page 5-6; 115pp; English.

XX CC The present sequence encodes a human metalloprotease enzyme designated

XX CC IGS5. IGS5 polynucleotides and polypeptides are useful for treating

XX CC infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma,

XX CC Parkinson's disease, acute heart failure, hypotension, hypertension,

XX CC urinary retention, osteoporosis, angina pectoris, myocardial infarction,

XX CC stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,

XX CC psychotic and neurological disorders, autism, multiple sclerosis,

XX CC Alzheimer's disease, and other neurodegenerative diseases, sleep

XX CC disorder, epilepsy, cardiovascular diseases, arteriosclerosis,

XX CC cerebrovasospasm, subarachnoid hemorrhage, cerebral ischemia, cerebral

XX CC infarction, peripheral vascular disease, Raynaud's disease, kidney

XX CC diseases, gastrointestinal disorders, motility disorders and conditions

XX CC of delayed gastric emptying, post-operative or diabetic gastroparesis,

XX CC diarrhoea, inflammations, chemotherapy induced injury, tumour invasion,

XX CC immune disorders, arthritis, endotoxin shock, sepsis, complications of

XX CC diabetes mellitus, and severe mental retardation and dyskinesias, such

XX CC as Huntington's disease or Gilles de la Tourette's syndrome.

XX SQ Sequence 2076 BP; 493 A; 578 C; 636 G; 369 T; 0 other;

Query Match 84.8%; Score 17.8; DB 22; Length 2076;  
Best Local Similarity 90.5%; Pred. No. 24;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCGAAGTTCTTGAGGCTCC 21  
||||||| ||| |||||  
Db 1405 CCCGAAGTTCTTGAGGCTCC 1385

## RESULT 8

AAS97186/c  
ID AAS97186 standard; cDNA; 2232 BP.

AC AAS97186;

DT 26-FEB-2002 (first entry)

DE Human metalloprotease partial DNA sequence #15.

XX Human; protease; PCR primer; cytostatic; immunomodulator; cardiant;  
KW vasotropic; antitumor; analgesic; endocrine; neurotrophic; tranquilizer;  
KW hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic;  
KW anorectic; antinflammatory; aspartyl protease; cysteine protease;  
KW metalloprotease; serine protease; cancer; haematopoietic; breast; colon;  
KW lung; prostate; cervical; brain; ovarian; bladder; kidney; pain;  
KW immune-related disease; cardiovascular disease; neuronal disease;  
KW migraine; sexual dysfunction; mood disorder; attention disorder;  
KW cognition disorder; hypotension; hypertension; psychotic disorder;  
KW dyskinesia; metabolic disorder; inflammatory disorder; ss.

XX Homo sapiens.

XX WO200183782-A2.

XX 08-NOV-2001.

XX 04-MAY-2001; 2001WO-US14431.

XX 04-MAY-2000; 2000US-201879P.

XX (SUGB-) SUGEN INC.

XX Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;  
Payne V;

XX WPI; 2002-041502/05.

XX P-PSDB; AAU72903.

XX Novel protease polypeptide useful for screening for substances that may  
be used to treat, e.g., cancers, immune-related diseases,  
cardiovascular disease, migraine, pain, psychotic and inflammatory  
disorders -

XX Claim 30; Figure 1R-S; 232pp; English.

XX The invention relates to an isolated, enriched, or purified protease  
polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to  
screen for substances (S) that may modulate its activity. Administering  
S (which modulates protease activity in vitro) may be used to treat a  
disease or disorder selected from cancers (e.g., of tissues, of blood or  
haematopoietic origin, of the breast, colon, lung, prostate, cervical,  
brain, ovarian, bladder or kidney), immune-related diseases and  
disorders, cardiovascular disease, brain or neuronal-associated diseases  
(e.g., central or peripheral nervous system diseases, migraine, pain,  
sexual dysfunction, mood disorders, attention disorders, cognition  
disorders, hypotension, hypertension, psychotic disorders, neurological  
disorders and dyskinesias), metabolic disorders and inflammatory  
disorders. (I) may also be useful as a diagnostic tool for a disease or  
disorder such as those above. AAS97159-AAS97195 represent human  
protease coding sequences and primers of the invention.

XX Sequence 2232 BP; 512 A; 620 C; 705 G; 395 T; 0 other;

Query Match 84.8%; Score 17.8; DB 24; Length 2232;  
Best Local Similarity 90.5%; Pred. No. 24;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCGAAGTTCTTGAGGCTCC 21  
||||||| ||| |||||  
Db 1561 CCCGAAGTTCTTGAGGCTCC 1541

## RESULT 9

AAF89739/c

ID AAF89739 standard; DNA; 2262 BP.

XX AAF89739;

XX 23-JUL-2001 (first entry)

XX Nucleotide sequence of a human metalloprotease enzyme IGS5.

XX Metalloprotease; IGS5; infection; pain; cancer; diabetes; obesity;  
KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;  
KW hypotension; hypertension; urinary retention; osteoporosis;  
KW angina pectoris; myocardial infarction; stroke; ulcer; allergy;  
KW benign prostatic hypertrophy; migraine; psychotic disorder;  
KW neurological disorder; autism; multiple sclerosis; Alzheimer's disease;  
KW neurodegenerative disease; sleep disorder; epilepsy; kidney disease;  
KW cardiovascular disease; arteriosclerosis; cerebrovasospasm;  
KW subarachnoid hemorrhage; cerebral ischemia; cerebral infarction;  
KW peripheral vascular disease; Raynaud's disease; motility disorder;  
KW gastrointestinal disorder; gastric emptying; gastroparesis; diarrhoea;  
KW inflammation; chemotherapy induced injury; tumour invasion;  
KW immune disorder; arthritis; endotoxin shock; sepsis; diabetes mellitus;  
KW severe mental retardation; dyskinesia; Huntington's disease;  
KW Gilles de la Tourette's syndrome; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..2262

XX /\*tag= a

XX /product= "metalloprotease enzyme IGS5"

XX WO200136610-A1.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-EP11532.

XX 19-NOV-1999; 99EP-0203862.

XX 19-NOV-1999; 99NL-1013616.

XX 31-MAY-2000; 2000EP-0201937.

XX 31-MAY-2000; 2000NL-1015356.

XX (SOLV ) SOLVAY PHARM BV.

XX Deleersnijder W, Wiegers R, Weske M;

XX WPI; 2001-343815/36.

XX P-PSDB; AAB83842.

XX New IGS5 polypeptides useful for treating infections, pain, cancer,  
diabetes, obesity, anorexia, bulimia, asthma, schizophrenia,  
hypertension, urinary retention and Parkinson's disease -

XX Claim 11; Page 8-9; 115pp; English.

XX The present sequence encodes a human metalloprotease enzyme designated  
IGS5. IGS5 polynucleotides and polypeptides are useful for treating  
infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma,  
Parkinson's disease, acute heart failure, hypotension, hypertension,  
urinary retention, osteoporosis, angina pectoris, myocardial infarction,  
stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,

psychotic and neurological disorders, autism, multiple sclerosis, Alzheimer's disease, and other neurodegenerative diseases, sleep disorder, epilepsy, cardiovascular diseases, arteriosclerosis, cerebrovaspasm, subarachnoid hemorrhage, cerebral ischemia, cerebral infarction, peripheral vascular disease, Raynaud's disease, kidney diseases, gastrointestinal disorders, motility disorders and conditions of delayed gastric emptying, post-operative or diabetic gastroparesis, diarrhoea, inflammations, chemotherapy induced injury, tumour invasion, immune disorders, arthritis, endotoxin shock, sepsis, complications of diabetes mellitus, and severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome.

XX  
SQ Sequence 2262 BP; 520 A; 628 C; 716 G; 398 T; 0 other;

Query Match 84.8%; Score 17.8; DB 22; Length 2262;  
Best Local Similarity 90.5%; Pred. No. 24;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CCGGAAGTTCTTGAGGCTCC 21  
||||||| ||| |||||||||  
Db 1591 CCGGAAGTTCTTGAGGCTCC 1571

RESULT 10  
AAD30580/c  
ID AAD30580 standard; cDNA; 2318 BP.  
XX  
AC AAD30580;  
XX  
DT 21-MAY-2002 (first entry)  
XX  
DE Human protease, PRTS-13 CDNA.  
XX  
KW Human; protease; PRTS-13; enzyme; gastritis; cirrhosis; Crohn's disease; gastrointestinal disorder; autoimmune; inflammatory; cell proliferative; cardiovascular; developmental; epithelial; neurological; reproductive; AIDS; Acquired Immune Deficiency Syndrome; allergy; rheumatoid arthritis; anaemia; asthma; atherosclerosis; hypertension; myocardial infarction; hepatitis; cancer; psoriasis; Cushing's syndrome; hypothyroidism; eczema; epilepsy; Alzheimer's disease; Huntington's disease; Parkinson's disease; Pick's disease; infertility; vitiligo; drug screening; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 338..1651  
FT /\*tag= a  
FT /product= "Human PRTS-13 protein"  
FT sig\_peptide 338..427  
FT /\*tag= b  
FT mat\_peptide 428..1648  
FT /\*tag= c  
FT /product= "Mature PRTS-13 protein"  
XX  
PN W0200208396-A2.  
XX  
PD 31-JAN-2002.  
XX  
PF 17-JUL-2001; 2001WO-US22397.  
XX  
PR 21-JUL-2000; 2000US-220063P.  
PR 28-JUL-2000; 2000US-221680P.  
PR 04-AUG-2000; 2000US-223544P.  
PR 11-AUG-2000; 2000US-224717P.  
PR 16-AUG-2000; 2000US-225988P.  
PR 23-AUG-2000; 2000US-227568P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Deleageane AM, Gandhi AR, Hafalia AJA, Lu DAM, Patterson C;  
PI Tribouley CM, Das D, Kallick DA, Nguyen DB, Lee EA, Khan FA;  
PI Yue H, Au-Young J, Griffin JA, Policky JL, Ramkumar J, Yang J;  
PI Thangavelu K, Ding L, Kearney L, Baughn MR, Borowsky ML;

PI Sanjanwala MS, Yao MG, Burford N, Walla NK, Lal P, Lee S, Todd S;  
PI LO TP, Tang YT, Elliott VS, Azimzai Y, Lu Y;  
XX  
DR WPI: 2002-206082/26.  
DR P-PSDB: AAE19176.  
XX  
PT New human protease polypeptide, useful in diagnosis, prevention and treatment of gastrointestinal, cardiovascular, autoimmune/inflammatory, cell proliferative, developmental, epithelial and neurological disorders  
XX  
PS Claim 5; Page 174-175; 182pp; English.  
XX  
CC The invention relates to an isolated human protease polypeptide (PRTS). PRTS protein and DNA are useful for diagnosing, treating and preventing gastrointestinal disorders (gastritis, cirrhosis, Crohn's disease), autoimmune/inflammatory disorders (AIDS, allergy, rheumatoid arthritis, anaemia, asthma), cardiovascular disorder (atherosclerosis, hypertension, myocardial infarction), cell proliferative disorders (hepatitis, cancer, psoriasis), developmental disorders (Cushing's syndrome, hypothyroidism), epithelial disorder (vitiligo, keloid, eczema), neurological disorders (epilepsy, Alzheimer's disease, Pick's disease, Huntington's disease, Parkinson's disease), and reproductive disorders (infertility). PRTS protein is useful in a number of drug screening techniques and to analyse the proteome of a tissue or cell type. PRTS DNA is useful for creating knockin humanised animals or transgenic animals to model human diseases, in somatic or germline gene therapy and in microarrays utilising fluids or tissues from patients to detect altered PKIN expression. The present sequence is human PRTS-13 CDNA.  
XX  
SQ Sequence 2318 BP; 526 A; 647 C; 735 G; 410 T; 0 other;

Query Match 84.8%; Score 17.8; DB 24; Length 2318;  
Best Local Similarity 90.5%; Pred. No. 24;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CCGGAAGTTCTTGAGGCTCC 21  
||||||| ||| |||||||||  
Db 1650 CCGGAAGTTCTTGAGGCTCC 1630

RESULT 11  
AAF89738/c  
ID AAF89738 standard; DNA; 2340 BP.  
XX  
AC AAF89738;  
XX  
DT 23-JUL-2001 (first entry)  
XX  
DE Nucleotide sequence of a human metalloprotease enzyme IGS5.  
XX  
KW Metalloprotease; IGS5; infection; pain; cancer; diabetes; obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypertension; myocardial infarction; stroke; ulcer; allergy; benign prostatic hypertrophy; migraine; psychotic disorder; neurological disorder; autism; multiple sclerosis; Alzheimer's disease; neurodegenerative disease; sleep disorder; epilepsy; kidney disease; cardiovascular disease; arteriosclerosis; cerebrovaspasm; subarachnoid hemorrhage; cerebral ischemia; cerebral infarction; peripheral vascular disease; Raynaud's disease; motility disorder; gastrointestinal disorder; gastric emptying; gastroparesis; diarrhoea; inflammation; chemotherapy induced injury; tumour invasion; immune disorder; arthritis; endotoxin shock; sepsis; diabetes mellitus; severe mental retardation; dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2340  
FT /\*tag= a  
FT /product= "metalloprotease enzyme IGS5"

XX PN WO200136610-A1.  
 XX PD 25-MAY-2001.  
 XX PF 17-NOV-2000; 2000WO-EP11532.  
 XX PR 19-NOV-1999; 99EP-0203862.  
 XX PR 19-NOV-1999; 99NL-1013616.  
 XX PR 31-MAY-2000; 2000EP-0201937.  
 XX PR 31-MAY-2000; 2000NL-1015356.  
 XX (SOLV ) SOLVAY PHARM BV.  
 XX PA Deleersnijder W, Wiegers R, Weske M;  
 XX PI WPI: 2001-343815/36.  
 XX DR P-PSDB; AAB83841.  
 XX New IGS5 polypeptides useful for treating infections, pain, cancer,  
 PT diabetes, obesity, anorexia, bulimia, asthma, schizophrenia,  
 PT hypertension, urinary retention and Parkinson's disease.  
 XX Claim 11; Page 6-7; 115pp; English.  
 XX The present sequence encodes a human metalloprotease enzyme designated  
 CC IGS5. IGS5 polynucleotides and polypeptides are useful for treating  
 CC infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma,  
 CC Parkinson's disease, acute heart failure, hypotension, hypertension,  
 CC urinary retention, osteoporosis, angina pectoris, myocardial infarction,  
 CC stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,  
 CC psychotic and neurological disorders, autism, multiple sclerosis,  
 CC Alzheimer's disease, and other neurodegenerative diseases, sleep  
 CC disorder, epilepsy, cardiovascular diseases, arteriosclerosis,  
 CC cerebrovasospasm, subarachnoid hemorrhage, cerebral ischemia, cerebral  
 CC infarction, peripheral vascular disease, Raynaud's disease, kidney  
 CC diseases, gastrointestinal disorders, motility disorders and conditions  
 CC of delayed gastric emptying, post-operative or diabetic gastroparesis,  
 CC diarrhoea, inflammations, chemotherapy induced injury, tumour invasion,  
 CC immune disorders, arthritis, endotoxin shock, sepsis, complications of  
 CC diabetes mellitus, and severe mental retardation and dyskinesias, such  
 CC as Huntington's disease or Gilles de la Tourette's syndrome.  
 XX Sequence 2340 BP; 539 A; 649 C; 739 G; 413 T; 0 other;  
 SQ  
 Query Match 84.8%; Score 17.8; DB 22; Length 2340;  
 Best Local Similarity 90.5%; Pred. No. 24;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CCCGAAGTTCTTGAGGCTCC 21  
 Db 1669 CCCGAAGTTCTTGAGGCTCC 1649  
 RESULT 12  
 ABN84280/C  
 ID ABN84280 standard; cDNA: 2580 BP.  
 XX AC ABN84280;  
 XX 23-SEP-2002 (first entry)  
 XX Human SEP endopeptidase coding sequence.  
 XX SEP; endopeptidase; enzyme; human; neuropeptide Y1; receptor;  
 KW male sexual dysfunction; male erectile dysfunction; obesity;  
 KW anorexia; bulimia; vasotropic; anorectic; therapy; inhibitor; gene;  
 KW ss.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 XX misc\_feature 1..65

FT /\*tag= a  
 FT /\*note= "5' partial vector sequence"  
 FT 258..2580  
 FT /\*tag= b  
 FT /\*product= "SEP"  
 FT /\*partial  
 FT /\*note= "the CDS does not include a stop codon"  
 FT 2580  
 FT /\*tag= c  
 FT /\*note= "the 3' end of the sequence is missing  
 FT from Figure 8 of the specification."  
 XX  
 XX PN WO200247670-A1.  
 XX PD 20-JUN-2002.  
 XX PF 10-DEC-2001; 2001WO-IB02399.  
 XX PR 15-DEC-2000; 2000GB-0030647.  
 XX PR 06-APR-2001; 2001GB-0008730.  
 XX PR 23-APR-2001; 2001GB-0009910.  
 XX PR 04-MAY-2001; 2001GB-0011037.  
 XX PR 29-JUN-2001; 2001US-0895367.  
 XX PR 13-JUL-2001; 2001US-0905846.  
 XX PR 24-AUG-2001; 2001GB-0020679.  
 XX (PFIZ ) PFIZER LTD.  
 XX (PFIZ ) PFIZER INC.  
 XX Naylor AM, Van Der Graaf PH, Wayman CP;  
 WPI: 2002-547828/58.  
 P-PSDB; ABB79521.  
 Use of an inhibitor of neuropeptide Y in the preparation of medicament  
 for the treatment or prevention of male erectile dysfunction  
 Disclosure; Fig 8; 179pp; English.  
 The present sequence is a nucleotide sequence coding for human  
 CC SEP, a soluble secreted endopeptidase, with an additional 5'  
 CC vector partial sequence. The invention relates to the use of an  
 CC inhibitor of NPY, or an inhibitor of NPY Y1 receptor; especially an  
 CC inhibitor selective for an NPY or NPY Y1 receptor associated with  
 CC male genitalia, in the preparation of a medicament for the  
 CC treatment or prevention of male sexual dysfunction, especially male  
 CC erectile dysfunction (MED). The NPY inhibitor may be used with an  
 CC auxiliary active agent such as an SEP inhibitor. The invention  
 CC provides a SEP assay that can be used to detect candidate  
 CC inhibitors of SEP. In addition to treatment of MED, NPY inhibitors  
 CC can also be used to treat abnormal drink and food intake disorders,  
 CC such as obesity, bulimia, anorexia and metabolic disorders.  
 XX Sequence 2580 BP; 587 A; 741 C; 794 G; 458 T; 0 other;  
 SQ  
 Query Match 84.8%; Score 17.8; DB 24; Length 2580;  
 Best Local Similarity 90.5%; Pred. No. 25;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CCCGAAGTTCTTGAGGCTCC 21  
 Db 1918 CCCGAAGTTCTTGAGGCTCC 1898  
 RESULT 13  
 AAF59660/C  
 ID AAF59660 standard; cDNA: 2636 BP.  
 XX AC AAF59660;  
 XX 27-APR-2001 (first entry)  
 XX Human neprilysin-like membrane metalloproteinase SNEPb cDNA.

XX Human; SNEPb; neprilysin-like membrane metalloproteinase;  
 KW splice variant; alternative splicing; zinc endopeptidase family;  
 KW neuropeptide; peptide hormone; processing; metabolism; vaccine;  
 KW drug screening; renal insufficiency; hepatic disease; ischaemia; pain;  
 KW hypertension; cancer; inflammation; cardiovascular disease;  
 KW neuronal disease; pancreatic disease; prostatic disease;  
 KW respiratory disease; gene therapy; nephrotropic; vasotropic; analgesic;  
 KW cerebroprotective; hypotensive; cytostatic; antiinflammatory;  
 KW cardiovascular; hepatotropic; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 XX EP1069188-A1.  
 PN  
 XX  
 XX 17-JAN-2001.  
 PD  
 XX  
 XX 15-JUL-1999; 99EP-0401767.  
 PF  
 XX  
 XX 15-JUL-1999; 99EP-0401767.  
 PR  
 XX  
 XX (SNFI ) SANOFI-SYNTHELABO.  
 PA  
 XX  
 XX Jagerschmidt A, Agnel M, Culouscou J;  
 PI  
 XX  
 XX WPI; 2001-212582/22.  
 DR  
 XX  
 XX P-PSDB; AAB60562.  
 DR  
 XX  
 XX New membrane-associated metalloproteinase SNEPa, SNEPb and SNEPc  
 PT polypeptides and polynucleotides, useful for treating e.g. acute and  
 PT chronic renal insufficiency, pain, stroke, cancer, inflammation, renal  
 PT and hepatic ischemia  
 PT  
 XX  
 XX Claim 5; Page 30-33; 72pp; English.  
 PS  
 XX  
 XX The invention relates to the human SNEPa, SNEPb and SNEPc proteins, and  
 CC the cDNAs encoding them. SNEPa, SNEPb and SNEPc are neprilysin-like  
 CC membrane metalloproteinases and are the products of alternative splicing.  
 CC The substrate(s) for the SNEP proteins are not as yet known, although  
 CC the neprilysin family of zinc endopeptidases play key roles in the  
 CC processing and/or metabolism of neuropeptides and peptide hormones. SNEP  
 CC nucleotides may be used as hybridisation probes for cDNA and genomic  
 CC DNA; to isolate full-length cDNAs and genomic clones encoding SNEPa,  
 CC SNEPb or SNEPc; to isolate cDNA and genomic clones of SNEP homologues;  
 CC as research reagents and material for the discovery of treatments and  
 CC diagnostics for animal and human diseases; and for chromosome  
 CC identification. The SNEP proteins may be used as immunogens to  
 CC produce antibodies immunospecific for SNEPa, SNEPb or SNEPc. Such  
 CC antibodies are used to isolate or identify clones expressing the  
 CC protein, or to purify the proteins by affinity chromatography.  
 CC SNEP proteins may also be used in screening for compounds which modulate  
 CC SNEP endopeptidase activity, and to assess enzymatic cleavage of small  
 CC molecule substrates in cells, cell-free preparations, chemical libraries  
 CC and product mixtures. The SNEP proteins (as vaccine compositions),  
 CC SNEP nucleotides, and SNEP activators or inhibitors may be used  
 CC to treat acute and chronic renal insufficiency, renal and hepatic  
 CC ischaemia, pain, stroke, hypertensive disease, cancer, inflammation, as  
 CC well as cardiovascular, neuronal, pancreatic, prostatic, renal,  
 CC respiratory or hepatic diseases. They may also be used in modulating  
 CC peptide activation and/or degradation in the brain or kidney or in  
 CC another organ, or to diagnose or treat any disorder related to abnormal  
 CC expression of SNEPa, SNEPb or SNEPc. The present sequence represents  
 CC cDNA encoding SNEPb.  
 XX  
 XX Sequence 2636 BP; 601 A; 759 C; 813 G; 462 T; 1 other;  
 SQ  
 Query Match 84.8%; Score 17.8; DB 22; Length 2636;  
 Best Local Similarity 90.5%; Pred. No. 25;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CCCGAAGTTTCTTGTAGGCTCC 21  
 ||||| ||| ||||| |||||  
 DB 1598 CCCGAGGTTCTGTAGGCTCC 1578

## RESULT 14

AAF59661/C

ID AAF59661 standard; cDNA; 2663 BP.

XX

AC AAF59661;

XX

DT 27-APR-2001 (first entry)

XX

DE Human neprilysin-like membrane metalloproteinase SNEPc cDNA.

XX

XX

KW Human; SNEPc; neprilysin-like membrane metalloproteinase;

KW

KW splice variant; alternative splicing; zinc endopeptidase family;

KW

KW neuropeptide; peptide hormone; processing; metabolism; vaccine;

KW

KW drug screening; renal insufficiency; hepatic disease; ischaemia; pain;

KW

KW hypertension; cancer; inflammation; cardiovascular disease;

KW

KW neuronal disease; pancreatic disease; prostatic disease;

KW

KW respiratory disease; gene therapy; nephrotropic; vasotropic; analgesic;

KW

KW cerebroprotective; hypotensive; cytostatic; antiinflammatory;

KW

KW cardiovascular; hepatotropic; ss.

XX

XX

OS Homo sapiens.

XX

XX

EP1069188-A1.

PN

XX

PD 17-JAN-2001.

XX

PF 15-JUL-1999; 99EP-0401767.

XX

PR 15-JUL-1999; 99EP-0401767.

XX

(SNFI ) SANOFI-SYNTHELABO.

PA

XX

PI Jagerschmidt A, Agnel M, Culouscou J;

XX

WPI; 2001-212582/22.

DR

P-PSDB; AAB60563.

XX

PT New membrane-associated metalloproteinase SNEPa, SNEPb and SNEPc

PT

PT polypeptides and polynucleotides, useful for treating e.g. acute and

PT

PT chronic renal insufficiency, pain, stroke, cancer, inflammation, renal

PT

PT and hepatic ischemia

XX

PS Claim 5; Page 35-38; 72pp; English.

XX

CC The invention relates to the human SNEPa, SNEPb and SNEPc proteins, and

CC

CC the cDNAs encoding them. SNEPa, SNEPb and SNEPc are neprilysin-like

CC

CC membrane metalloproteinases and are the products of alternative splicing.

CC

CC The substrate(s) for the SNEP proteins are not as yet known, although

CC

CC the neprilysin family of zinc endopeptidases play key roles in the

CC

CC processing and/or metabolism of neuropeptides and peptide hormones. SNEP

CC

CC nucleotides may be used as hybridisation probes for cDNA and genomic

CC

CC DNA; to isolate full-length cDNAs and genomic clones encoding SNEPa,

CC

CC SNEPb or SNEPc; to isolate cDNA and genomic clones of SNEP homologues;

CC

CC as research reagents and material for the discovery of treatments and

CC

CC diagnostics for animal and human diseases; and for chromosome

CC

CC identification. The SNEP proteins may be used as immunogens to

CC

CC produce antibodies immunospecific for SNEPa, SNEPb or SNEPc. Such

CC

CC antibodies are used to isolate or identify clones expressing the

CC

CC protein, or to purify the proteins by affinity chromatography.

CC

CC SNEP proteins may also be used in screening for compounds which modulate

CC

CC SNEP endopeptidase activity, and to assess enzymatic cleavage of small

CC

CC molecule substrates in cells, cell-free preparations, chemical libraries,

CC

CC and product mixtures. The SNEP proteins (as vaccine compositions),

CC

CC SNEP nucleotides, and SNEP activators or inhibitors may be used

CC

CC to treat acute and chronic renal insufficiency, renal and hepatic

CC

CC ischaemia, pain, stroke, hypertensive disease, cancer, inflammation, as

CC

CC well as cardiovascular, neuronal, pancreatic, prostatic, renal,

CC

CC respiratory or hepatic diseases. They may also be used in modulating

CC

CC peptide activation and/or degradation in the brain or kidney or in

CC

CC another organ, or to diagnose or treat any disorder related to abnormal

CC

CC expression of SNEPa, SNEPb or SNEPc. The present sequence represents

CC

## CC CDNA encoding SNEPC:

XX Sequence 2663 BP; 609 A; 769 C; 816 G; 468 T; 1 other;

Query Match 84.8%; Score 17.8; DB 22; Length 2663;  
 Best Local Similarity 90.5%; Pred. No. 25;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCGAAGTTTCTTGAGGCTCC 21  
 ||||| ||| |||||  
 Db 1625 CCCGAAGCTTCTTGAGGCTCC 1605

## RESULT 15

AAA63764/C

ID AAA63764 standard; cDNA: 2676 BP;

XX AC AAA63764;

DT 04-DEC-2000 (first entry)

XX cDNA encoding neutral endopeptidase metalloproteinase-like enzyme NL-2.

XX Neprilysin; neutral endopeptidase metalloproteinase-like enzyme;  
 KW NEP-like enzyme; protein production; protein secretion;  
 KW neurological disease; Alzheimer's disease; pain; psychiatric disorder;  
 KW fertility; bone disease; abnormal phosphate metabolism; ss.

XX Homo sapiens.

FH Key Location/Qualifiers  
 FT CDS 7..2319  
 FT /\*tag= a  
 FT /product= "neutral endopeptidase metalloproteinase-like  
 FT enzyme NL-2"

PN WO200047750-A2.

XX 17-AUG-2000.

PF 11-FEB-2000; 2000WO-CA00147.

XX 11-FEB-1999; 99CA-2260376.

XX (UIMO-) UNIV MONTREAL.

XX Desgroseillers L, Boileau G;

XX WPI; 2000-549148/50.

DR P-PSDB; AAB08131.

XX Novel neutral endopeptidase-like metalloproteinase polypeptides and  
 PT polynucleotides, used to screen for related sequences and enzyme  
 PT inhibitors, used for the treatment of NL-3 related bone disorders -

PS Disclosure; Fig 4; 59pp; English.

XX The present sequence encodes a human neutral endopeptidase  
 CC metalloproteinase-like enzyme, designated NL-2. The specification  
 CC also describes NL-1 and NL-3. The NL enzymes are used to test for  
 CC specific inhibitors. The N-terminal region of the enzymes can be used  
 CC to promote production and secretion of foreign proteins and active  
 CC biopeptides, using chimeric constructs containing the foreign protein  
 CC downstream from and in phase with the N-terminal region. The NL enzymes  
 CC are have been localised to the brain, and may be useful in the  
 CC treatment of neurological diseases such as Alzheimer's disease, pain,  
 CC and psychiatric disorders. NL enzymes have also been localised to the  
 CC testis and ovaries, and may be used to control fertility. They have  
 CC also been localised to bones, and may be used to treat bone diseases,  
 CC and abnormal phosphate metabolisms related to improper peptide  
 CC processing by the NL-3 enzyme.

XX Sequence 2676 BP; 608 A; 771 C; 823 G; 474 T; 0 other;

Query Match 84.8%; Score 17.8; DB 21; Length 2676;  
 Best Local Similarity 90.5%; Pred. No. 25;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCGAAGTTTCTTGAGGCTCC 21  
 ||||| ||| |||||  
 Db 1648 CCCGAAGCTTCTTGAGGCTCC 1628

Search completed: July 8, 2003, 02:19:12  
 Job time : 133.941 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 01:24:03 ; Search time 28.9902 Seconds  
(without alignments)  
222.151 Million cell updates/sec

Title: US-09-647-780A-18  
Perfect score: 21  
Sequence: 1 cccgaagtcttcttgagctcc 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	16.2	77.1	199	3	US-08-961-083-91
C 2	16.2	77.1	3072	4	US-09-522-217-55
C 3	16.2	77.1	6201	2	US-08-790-912-1
C 4	16.2	77.1	15213	4	US-08-961-527-26
C 5	15.2	72.4	30	3	US-08-961-083-317
C 6	15.2	72.4	889	1	US-08-832-883-52
C 7	15.2	72.4	889	2	US-08-832-877-52
C 8	15.2	72.4	1386	4	US-09-134-001C-1055
C 9	15.2	72.4	3914	1	US-08-117-373-11
C 10	15.2	72.4	6924	1	US-08-015-973-2
C 11	15.2	72.4	6924	2	US-08-448-164-2
C 12	15.2	72.4	7941	4	US-09-816-703A-1
C 13	15.2	72.4	4403765	4	US-09-103-840A-2
C 14	15.2	72.4	4411529	4	US-09-103-840A-1
C 15	14.8	70.5	685	4	US-09-118-442-22
C 16	14.8	70.5	685	4	US-09-677-064-22
C 17	14.8	70.5	875	2	US-08-867-676-2
C 18	14.8	70.5	1621	4	US-09-227-357-29
C 19	14.8	70.5	1782	4	US-09-537-877-11
C 20	14.8	70.5	2991	1	US-08-324-977-49
C 21	14.8	70.5	2991	2	US-08-384-616-49
C 22	14.8	70.5	2991	2	US-08-904-686A-49
C 23	14.8	70.5	2991	4	US-09-315-850-49
C 24	14.8	70.5	3507	1	US-08-315-468-3
C 25	14.8	70.5	6642	1	US-08-727-034-5
C 26	14.8	70.5	6843	1	US-08-727-034-6
C 27	14.8	70.5	7863	1	US-08-324-977-35

C 28	14.8	70.5	7863	2	US-08-384-616-35	Sequence 35, Appl
C 29	14.8	70.5	7863	2	US-08-904-686A-35	Sequence 35, Appl
C 30	14.8	70.5	7863	4	US-09-315-850-35	Sequence 35, Appl
C 31	14.8	70.5	7917	1	US-08-324-977-31	Sequence 31, Appl
C 32	14.8	70.5	7917	2	US-08-384-616-31	Sequence 31, Appl
C 33	14.8	70.5	7917	2	US-08-904-686A-31	Sequence 31, Appl
C 34	14.8	70.5	7917	4	US-09-315-850-31	Sequence 31, Appl
C 35	14.8	70.5	9030	1	US-08-324-977-13	Sequence 13, Appl
C 36	14.8	70.5	9030	2	US-08-384-616-13	Sequence 13, Appl
C 37	14.8	70.5	9030	2	US-08-904-686A-13	Sequence 13, Appl
C 38	14.8	70.5	9030	4	US-09-315-850-13	Sequence 13, Appl
C 39	14.8	70.5	9416	1	US-08-324-977-1	Sequence 1, Appl
C 40	14.8	70.5	9416	2	US-08-384-616-1	Sequence 1, Appl
C 41	14.8	70.5	9416	2	US-08-904-686A-1	Sequence 1, Appl
C 42	14.8	70.5	9416	4	US-09-315-850-1	Sequence 1, Appl
C 43	14.8	70.5	9416	4	US-08-823-895A-27	Sequence 27, Appl
C 44	14.8	70.5	9595	3	US-09-014-416-4	Sequence 4, Appl
C 45	14.6	69.5	825	3	US-09-120-582-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-08-961-083-91/C  
; Sequence 91, Application US/08961083  
; Patent No. 6159469  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961.083  
; FILING DATE:  
; CLASSIFICATION: 435.  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 91:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 199 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-961-083-91

Query Match 77.1%; Score 16.2; DB 3; Length 199;  
Best Local Similarity 85.7%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCCGAAGTTCTTGAGGCTCC 21  
DB 21 CCCGAGTTCTTGAGCATCC 1



RESULT 2  
US-09-522-217-55  
Sequence 55, Application US/09522217  
Patent No. 6307024  
GENERAL INFORMATION:  
APPLICANT: No. 6307024ak, Julia E.  
APPLICANT: Presnell, Scott R.  
APPLICANT: Sprecher, Cindy A.  
APPLICANT: Foster, Donald C.  
APPLICANT: Holly, Richard D.  
APPLICANT: Gross, Jane A.  
APPLICANT: Johnston, Janet V.  
APPLICANT: Nelson, Andrew J.  
APPLICANT: Dillon, Stacey R.  
APPLICANT: Hammond, Angela K.  
TITLE OF INVENTION: NOVEL CYTOKINE ZALPHAII LIGAND  
FILE REFERENCE: 99-16  
CURRENT APPLICATION NUMBER: US/09/522,217  
CURRENT FILING DATE: 2000-03-09  
EARLIER FILING DATE: 1999-03-09  
EARLIER APPLICATION NUMBER: US 60/123,547  
EARLIER FILING DATE: 1999-03-09  
EARLIER APPLICATION NUMBER: US 60/123,904  
EARLIER FILING DATE: 1999-03-11  
EARLIER APPLICATION NUMBER: US 60/142,013  
EARLIER FILING DATE: 1999-07-01  
NUMBER OF SEQ ID NOS: 115  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 55  
LENGTH: 3072  
TYPE: DNA  
ORGANISM: mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (54)...(491)  
US-09-522-217-55

Query Match 77.1%; Score 16.2; DB 4; Length 3072;  
Best Local Similarity 85.7%; Pred. No. 20;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 CCGAAGTTCTTGTGAGGCTCC 21  
||||| ||| ||||| |||  
Db 506 CCGAAGATTCCTGAGCATCC 526  
RESULT 3  
US-08-790-912-1/c  
Sequence 1, Application US/08790912  
Patent No. 5976542  
GENERAL INFORMATION:  
APPLICANT: Weiser, Jeffrey N.  
APPLICANT: Plaut, Andrew G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT  
OF STREPTOCOCCUS PNEUMONIAE INFECTION  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.  
STREET: 1601 Market Street, 36th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103-2398  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/790,912  
FILING DATE: 29-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/026,939  
FILING DATE: 23-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Leary, Kathryn  
REGISTRATION NUMBER: 36,317  
REFERENCE/DOCKET NUMBER: 7600-401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 567-2020  
TELEFAX: (215) 567-2991  
TELEX: 831-494  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6201 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..6201  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(1..30, 34..39, 46..51, 55..99, 103..135, 139  
153, 157..165, 169..186, 190..210, 217..225,  
LOCATION: 229..234, 238..6150, 6154..6177, 6181..6201)  
US-08-790-912-1  
Query Match 77.1%; Score 16.2; DB 2; Length 6201;  
Best Local Similarity 85.7%; Pred. No. 22;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 CCGAAGTTCTTGTGAGGCTCC 21  
||||| ||| ||||| |||  
Db 398 CCGCAGTTCTTGTGAGCATCC 378  
RESULT 4  
US-08-961-527-26  
Sequence 26, Application US/08961527  
Patent No. 6420135  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 391  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,527  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:  
LENGTH: 15213 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-961-527-26

Query Match 77.1%; Score 16.2; DB 4; Length 15213;  
Best Local Similarity 85.7%; Pred. No. 25;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGAAGTTCTTGAGGCTCC 21  
DB 5337 CCGCAGTTCTTGAGCATCC 5357

RESULT 5  
US-08-961-083-317/c  
Sequence 317, Application US/08961083  
Patent No. 6159469  
GENERAL INFORMATION:  
APPLICANT: Choi et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,083  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P2  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 317:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-961-083-317

Query Match 72.4%; Score 15.2; DB 3; Length 30;  
Best Local Similarity 85.0%; Pred. No. 32;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCGAAGTTCTTGAGGCTCC 21  
DB 30 CCGCAGTTCTTGAGCATCC 11

RESULT 6  
US-08-832-883-52  
Sequence 52, Application US/08832883  
Patent No. 5807681  
GENERAL INFORMATION:  
APPLICANT: Giordano, Antonio

APPLICANT: Baldi, Alphonso  
TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS  
OF CANCER  
NUMBER OF SEQUENCES: 115  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.  
STREET: Suite 1800 Two Penn Center Plaza  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19102

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/832,883  
FILING DATE:

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 8321-13 US1  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 889 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-832-883-52

Query Match 72.4%; Score 15.2; DB 1; Length 889;  
Best Local Similarity 85.0%; Pred. No. 54;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCGAAGTTCTTGAGGCTCC 21  
DB 191 CCGAAGTTCTTGAGCTTC 210

RESULT 7  
US-08-832-877-52  
Sequence 52, Application US/08832877  
Patent No. 5840506  
GENERAL INFORMATION:  
APPLICANT: Giordano, Antonio  
TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS OF  
CANCER  
NUMBER OF SEQUENCES: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.  
STREET: Suite 1800 Two Penn Center Plaza  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19102

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/832,877  
FILING DATE:

CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A  
REGISTRATION NUMBER: 30,480

; REFERENCE/DOCKET NUMBER: 8321-13 US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; INFORMATION FOR SEQ ID NO: 52:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 889 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-832-877-52

Query Match 72.4%; Score 15.2; DB 2; Length 889;  
Best Local Similarity 85.0%; Pred. No. 54;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCGAAGTTCTTGAGGCTCC 21  
|||||  
Db 191 CCGAAGTTCTTAAAGCTTC 210

RESULT 8  
US-09-134-001C-1055/c  
; Sequence 1055, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: EPIDERMIS FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 1055  
; LENGTH: 1386  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-1055

Query Match 72.4%; Score 15.2; DB 4; Length 1386;  
Best Local Similarity 85.0%; Pred. No. 58;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCCGAAGTTCTTGAGGCTC 20  
|||||  
Db 875 CCCCAAGTTACTTGAGGTC 856

RESULT 9  
US-08-117-373-11/c  
; Sequence 11, Application US/08117373  
; Patent No. H001903  
; GENERAL INFORMATION:

; APPLICANT: BOWDEN, PAUL E.  
; APPLICANT: LUO, XIAOCHUN  
; APPLICANT: WAWRZYNIAK, CYNTHIA J.  
; TITLE OF INVENTION: ISOLATED HAIR KERATIN GENES AND THEIR  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: THE PROCTER & GAMBLE COMPANY  
; STREET: 11810 EAST MIAMI RIVER ROAD  
; CITY: CINCINNATI  
; STATE: OHIO  
; COUNTRY: USA  
; ZIP: 45239-8707  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/117,373  
; FILING DATE:

; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CORSTANJE, BRAHM J.  
; REGISTRATION NUMBER: 34,804  
; REFERENCE/DOCKET NUMBER: 4518R2&  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (513) 627-2858  
; TELEFAX: (513) 627-0260  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3914 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-117-373-11

Query Match 72.4%; Score 15.2; DB 1; Length 3914;  
Best Local Similarity 85.0%; Pred. No. 68;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCGAAGTTCTTGAGGCTCC 21  
|||||  
Db 915 CAGAGTTCTTGAGGCTCC 896

## RESULT 10

US-08-015-973-2  
; Sequence 2, Application US/08015973  
; Patent No. 5604094  
; GENERAL INFORMATION:

; APPLICANT: Schlessinger, Joseph  
; TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PENNIE & EDMONDS  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/015,973  
; FILING DATE: 10-FEB-1993  
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 7683-021  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212)869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6924 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS



## ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1:  
; LENGTH: 4411529

; TYPE: DNA

; ORGANISM: Mycobacterium tuberculosis

; OTHER INFORMATION: H37Rv

US-09-103-840A-1

Query Match 72.4%; Score 15.2; DB 4; Length 4411529;  
Best Local Similarity 85.0%; Pred. No. 64;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CGAAGTTCTTGAGGCTCC 21

||||| ||||| |||||

Db 4210838 CGAAGTTCTCGATGCTCC 4210819

## RESULT 15

US-09-118-442-22/c

; Sequence 22, Application US/09118442B

; Patent No. 6197561

; GENERAL INFORMATION:

; APPLICANT: Martino-Catt, Susan J.

; APPLICANT: Wang, Hongyu

; APPLICANT: Beach, Larry R.

; APPLICANT: Wang, Xun

; APPLICANT: Bowen, Benjamin A.

; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in

; FILE REFERENCE: 0706

; CURRENT APPLICATION NUMBER: US/09/118,442B

; CURRENT FILING DATE: 1998-07-17

; EARLIER APPLICATION NUMBER: 60/055,446

; EARLIER FILING DATE: 1997-08-11

; EARLIER APPLICATION NUMBER: 60/055,526

; EARLIER FILING DATE: 1997-08-08

; EARLIER APPLICATION NUMBER: 60/053,944

; EARLIER FILING DATE: 1997-07-28

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 22

; LENGTH: 685

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(685)

; OTHER INFORMATION: n = A,T,C or G

US-09-118-442-22

Query Match 70.5%; Score 14.8; DB 4; Length 685;  
Best Local Similarity 88.9%; Pred. No. 84;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CGAAGTTCTTGAGGCTC 20

||||| ||||| |||||

Db 157 CAAAGTTCTTGAGGATC 140

Search completed: July 8, 2003, 09:32:29  
Job time : 41.0402 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 19:09:54 ; Search time 121.288 Seconds  
(without alignments)  
273.390 Million cell updates/sec

Title: US-09-647-780A-18  
Perfect score: 21  
Sequence: 1 ccggaagtttcttgaggctcc 21

Scoring table: IDENTITY\_NUC  
Gap 10.0 , Gapext 1.0

Searched: 1105431 seqs, 789497651 residues

Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

- Database : Published\_Applications\_NA:\*
- 1: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
  - 2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
  - 3: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
  - 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
  - 5: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
  - 6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
  - 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
  - 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
  - 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
  - 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
  - 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
  - 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
  - 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
  - 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	17.8	84.8	2893	9	US-10-017-273A-4
C 2	17.8	84.8	2893	10	US-09-905-846-1
C 3	17.8	84.8	2975	9	US-10-017-273A-5
C 4	17.8	84.8	2975	10	US-09-905-846-5
C 5	16.2	77.1	199	10	US-09-785-272-31
C 6	16.2	77.1	3072	9	US-09-918-995-24415
C 7	16.2	77.1	3072	10	US-10-295-723-55
C 8	16.2	77.1	3072	10	US-09-923-246-55
C 9	16.2	77.1	3072	10	US-09-825-561A-46
C 10	15.8	75.2	368004	10	US-09-949-654-3
C 11	15.4	73.3	171	10	US-09-770-696-344
C 12	15.4	73.3	1175	9	US-10-051-307-4
C 13	15.4	73.3	2979	9	US-09-938-842A-48
C 14	15.4	73.3	7497	10	US-09-960-253-175
C 15	15.4	73.3	7792	12	US-10-044-090-359
C 16	15.2	72.4	30	10	US-09-765-272-317
C 17	15.2	72.4	213	9	US-09-933-797-788
C 18	15.2	72.4	258	9	US-09-918-995-29893
C 19	15.2	72.4	301	10	US-09-878-574-13668

Sequence 33959, A	9	US-09-918-995-33959
Sequence 358, App	9	US-09-991-936-358
Sequence 378, App	9	US-09-991-936-378
Sequence 1894, A	9	US-09-918-995-1894
Sequence 1804, Ap	9	US-09-991-936-1804
Sequence 1905, Ap	9	US-09-991-936-1905
Sequence 20, Appl	9	US-09-746-783-20
Sequence 108, App	10	US-09-764-864-108
Sequence 1901, Ap	9	US-09-991-936-1901
Sequence 1903, Ap	9	US-09-991-936-1903
Sequence 209, App	9	US-09-822-846-209
Sequence 8518, Ap	9	US-09-764-891-8518
Sequence 5, Appl	9	US-09-983-000A-5
Sequence 1812, Ap	10	US-09-954-456-1812
Sequence 3, Appl	9	US-09-983-000A-3
Sequence 1673, Ap	10	US-09-764-864-1673
Sequence 10, Appl	9	US-09-953-611-10
Sequence 3, Appl	9	US-09-901-152-3
Sequence 1, Appl	9	US-09-939-964-1
Sequence 17459, A	187	10
Sequence 671, App	472	10
Sequence 1710, Ap	955	10
Sequence 168, App	2961	9
Sequence 155, App	2991	10
Sequence 5392, Ap	267	10
Sequence 3535, Ap	443	10

ALIGNMENTS

RESULT 1  
US-10-017-273A-4/c  
; Sequence 4, Application US/10017273A  
; Publication No. US20030119714A1  
; GENERAL INFORMATION:  
; APPLICANT: Pfizer Inc.  
; APPLICANT: Naylor, Alasdair M.  
; APPLICANT: Van Der Graaf, Pieter H  
; APPLICANT: Wayman, Christopher P.  
; TITLE OF INVENTION: Treatment of Male Sexual Dysfunction  
; FILE REFERENCE: PC22013  
; CURRENT APPLICATION NUMBER: US/10/017,273A  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/265,358  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: GB 0030647.2  
; PRIOR FILING DATE: 2000-12-15  
; PRIOR APPLICATION NUMBER: GB 0108730.3  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: GB 0120679.6  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 09/905,846  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/291,722  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: US 09/895,367  
; PRIOR FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 2893  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-017-273A-4

Query Match 84.88; Score 17.8; DB 9; Length 2893;  
Best Local Similarity 90.5%; Pred. No. 12;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CCGGAAGTTTCTTGAGGCTCC 21  
DB 1853 CCGGAAGTTTCTTGAGGCTCC 1833

RESULT 2  
US-09-905-846-1/c  
; Sequence 1, Application US/09905846  
; Patent No. US20020102707A1  
; GENERAL INFORMATION:  
; APPLICANT: Ian Dennis Harrow  
; APPLICANT: Peter Stacey  
; APPLICANT: Roderick Thomas Walsh  
; TITLE OF INVENTION: Compounds for the treatment of sexual dysfunction  
; FILE REFERENCE: PCS10926APME  
; CURRENT APPLICATION NUMBER: US/09/905,846  
; CURRENT FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 0017387.2  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/220,908  
; PRIOR FILING DATE: 2000-07-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2893  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-905-846-1

Query Match 84.8%; Score 17.8; DB 10; Length 2893;  
Best Local Similarity 90.5%; Pred. No. 12;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CCCGAAGTTCTTGAGGCTCC 21  
Db 1853 CCCGAAGTTCTTGAGGCTCC 1833

RESULT 3  
US-10-017-273A-5/c  
; Sequence 5, Application US/10017273A  
; Publication No. US20030119714A1  
; GENERAL INFORMATION:  
; APPLICANT: Pfizer Inc.  
; APPLICANT: Naylor, Alasdair M.  
; APPLICANT: Van Der Graaf, Pieter H  
; APPLICANT: Wayman, Christopher P.  
; TITLE OF INVENTION: Treatment of Male Sexual Dysfunction  
; FILE REFERENCE: PC22013  
; CURRENT APPLICATION NUMBER: US/10/017,273A  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/265,358  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: GB 0030647.2  
; PRIOR FILING DATE: 2000-12-15  
; PRIOR APPLICATION NUMBER: GB 0108730.3  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: GB 0120679.6  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 09/905,846  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/291,722  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: US 09/895,367  
; PRIOR FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 2975  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-017-273A-5

Query Match 84.8%; Score 17.8; DB 9; Length 2975;  
Best Local Similarity 90.5%; Pred. No. 12;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CCCGAAGTTCTTGAGGCTCC 21  
Db 1918 CCCGAAGTTCTTGAGGCTCC 1898

RESULT 4  
US-09-905-846-5/c  
; Sequence 5, Application US/09905846  
; Patent No. US20020102707A1  
; GENERAL INFORMATION:  
; APPLICANT: Ian Dennis Harrow  
; APPLICANT: Peter Stacey  
; APPLICANT: Roderick Thomas Walsh  
; TITLE OF INVENTION: Compounds for the treatment of sexual dysfunction  
; FILE REFERENCE: PCS10926APME  
; CURRENT APPLICATION NUMBER: US/09/905,846  
; CURRENT FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 0017387.2  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/220,908  
; PRIOR FILING DATE: 2000-07-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 2975  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-905-846-5

Query Match 84.8%; Score 17.8; DB 10; Length 2975;  
Best Local Similarity 90.5%; Pred. No. 12;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CCCGAAGTTCTTGAGGCTCC 21  
Db 1918 CCCGAAGTTCTTGAGGCTCC 1898

RESULT 5  
US-09-765-272-91/c  
; Sequence 91, Application US/09765272  
; Patent No. US20020061545A1  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/765,272  
; FILING DATE: 22-Jan-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/961,083  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 91:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 199 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 91:  
US-09-765-272-91

Query Match  
Best Local Similarity 77.1%; Score 16.2; DB 10; Length 199;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCCGAAGTTTCTTGAGGCTCC 21  
||||| ||||| ||||| ||||| |||||

DB 21 CCCGAGTTTCTTGAGCATCC 1

## RESULT 6

US-09-918-995-24415/c  
Sequence 24415, Application US/09918995  
Publication No. US20030073623A1

## GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

FILE REFERENCE: 20411-756

CURRENT APPLICATION NUMBER: US/09/918,995

CURRENT FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: US/09/235,076

PRIOR FILING DATE: 1999-01-20

NUMBER OF SEQ ID NOS: 38054

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 24415

LENGTH: 508

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (1)...(508)

OTHER INFORMATION: n = A,T,C or G

US-09-918-995-24415

Query Match  
Best Local Similarity 77.1%; Score 16.2; DB 9; Length 508;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCCGAAGTTTCTTGAGGCTCC 21  
|| || ||||| ||||| |||||

DB 80 CCAGAGTTTCTTGAGTCTCC 60

## RESULT 7

US-10-295-723-55

Sequence 55, Application US/10295723

Publication No. US2003012524A1

## GENERAL INFORMATION:

APPLICANT: No. US2003012524A1ak, Julia E.

APPLICANT: Presnell, Scott R.

APPLICANT: Sprecher, Cindy A.

APPLICANT: Foster, Donald C.

APPLICANT: Holly, Richard D.

APPLICANT: Gross, Jane A.

APPLICANT: Johnston, Janet V.

APPLICANT: Nelson, Andrew J.

APPLICANT: Dillon, Stacey R.

APPLICANT: Hammond, Angela K.

TITLE OF INVENTION: NOVEL CYTOKINE ZALPHALL LIGAND

FILE REFERENCE: 99-16

CURRENT APPLICATION NUMBER: US/10/295,723

CURRENT FILING DATE: 2002-11-15

PRIOR APPLICATION NUMBER: 09/522,217

PRIOR FILING DATE: 2000-03-09

PRIOR APPLICATION NUMBER: US 60/123,547  
PRIOR FILING DATE: 1999-03-09  
PRIOR APPLICATION NUMBER: US 60/123,904  
PRIOR FILING DATE: 1999-03-11  
PRIOR APPLICATION NUMBER: US 60/142,013  
PRIOR FILING DATE: 1999-07-01  
NUMBER OF SEQ ID NOS: 115  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 55  
LENGTH: 3072  
TYPE: DNA  
ORGANISM: mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (54)...(491)  
US-10-295-723-55

Query Match  
Best Local Similarity 77.1%; Score 16.2; DB 9; Length 3072;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCCGAAGTTTCTTGAGGCTCC 21  
||||| ||||| ||||| ||||| |||||

DB 506 CCCGAGATTCTTGAGGATCC 526

## RESULT 8

US-09-923-246-55

Sequence 55, Application US/09923246

Patent No. US20020128446A1

## GENERAL INFORMATION:

APPLICANT: No. US20020128446A1ak, Julia E.

APPLICANT: Presnell, Scott R.

APPLICANT: Sprecher, Cindy A.

APPLICANT: Foster, Donald C.

APPLICANT: Holly, Richard D.

APPLICANT: Gross, Jane A.

APPLICANT: Johnston, Janet V.

APPLICANT: Nelson, Andrew J.

APPLICANT: Dillon, Stacey R.

APPLICANT: Hammond, Angela K.

TITLE OF INVENTION: NOVEL CYTOKINE ZALPHALL LIGAND

FILE REFERENCE: 99-16

CURRENT APPLICATION NUMBER: US/09/923,246

CURRENT FILING DATE: 2001-08-03

PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 2000-03-09

PRIOR FILING DATE: EARLIER FILING DATE: 2000-03-09

PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-03-11

PRIOR FILING DATE: EARLIER FILING DATE: 1999-03-11

PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-07-01

PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-01

NUMBER OF SEQ ID NOS: 115

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 55

LENGTH: 3072

TYPE: DNA

ORGANISM: mus musculus

FEATURE:

NAME/KEY: CDS

LOCATION: (54)...(491)

US-09-923-246-55

Query Match  
Best Local Similarity 77.1%; Score 16.2; DB 10; Length 3072;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCCGAAGTTTCTTGAGGCTCC 21  
||||| ||||| ||||| ||||| |||||

DB 506 CCCGAGATTCTTGAGGATCC 526

## RESULT 9

US-09-825-561A-46



; Sequence 46, Application US/09825561A  
; Patent No. US2002013767A1  
; GENERAL INFORMATION:

; APPLICANT: Sprecher, Cindy A.  
; APPLICANT: No. US2002013767A1ak, Julia E.

; APPLICANT: West, James W.

; APPLICANT: Presnell, Scott R.

; APPLICANT: Holly, Richard D.

; APPLICANT: Nelson, Andrew J.

; TITLE OF INVENTION: SOLUBLE ZALPHALL CYTOKINE RECEPTORS

; FILE REFERENCE: 00-22

; CURRENT APPLICATION NUMBER: US/09/825,561A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/194,731

; PRIOR FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/222,121

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 86

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 46

; LENGTH: 3072

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (54)...(491)

; US-09-825-561A-46

Query Match 77.1%; Score 16.2; DB 10; Length 3072;

Best Local Similarity 85.7%; Pred. No. 82;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCGAAGTTCTTGAGGCTCC 21

||||||| ||| ||||| |||

Db 506 CCGAAGATTCCTGAGGATCC 526

RESULT 10

US-09-949-654-3/c

; Sequence 3, Application US/09949654

; Patent No. US2002012764A1

; GENERAL INFORMATION:

; APPLICANT: YAN, Chunhua et al.

; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,

; TITLE OF INVENTION: AND USES THEREOF

; FILE REFERENCE: CL000817

; CURRENT APPLICATION NUMBER: US/09/949,654

; CURRENT FILING DATE: 2001-09-12

; PRIOR APPLICATION NUMBER: 60/231,572

; PRIOR FILING DATE: 2000-09-11

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 368004

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(368004)

; OTHER INFORMATION: n = A, T, C or G

US-09-949-654-3

Query Match 75.2%; Score 15.8; DB 10; Length 368004;

Best Local Similarity 89.5%; Pred. No. 1.8e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCGAAGTTCTTGAGGCT 19

||||||| ||||| |||||

Db 192753 CCGAAGTTTATTGAGGCT 192735

RESULT 11

US-09-770-696-344

; Sequence 344, Application US/09770696

; Patent No. US20010044940A1

; GENERAL INFORMATION:

; APPLICANT: Goriach, Jorn

; APPLICANT: An, Yong-Qiang

; APPLICANT: Hamilton, Carol M.

; APPLICANT: Price, Jennifer L.

; APPLICANT: Raines, Tracy M.

; APPLICANT: Yu, Yang

; APPLICANT: Rameaka, Joshua G.

; APPLICANT: Page, Amy

; APPLICANT: Matthew, Abraham V.

; APPLICANT: Ledford, Brooke L.

; APPLICANT: Woessner, Jeffrey P.

; APPLICANT: Haas, William David

; APPLICANT: Garcia, Carlos A.

; APPLICANT: Krickler, Maja

; APPLICANT: Slader, Ted

; APPLICANT: Davis, Keith R.

; APPLICANT: Allen, Keith

; APPLICANT: Hoffman, Neil

; APPLICANT: Hurban, Patrick

; TITLE OF INVENTION: Expressed Sequences of Arabidopsis

; TITLE OF INVENTION: thaliana

; FILE REFERENCE: 2031US (PARA-020PRV)

; CURRENT APPLICATION NUMBER: US/09/770,696

; CURRENT FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: 60/178,278

; PRIOR FILING DATE: 2000-01-27

; NUMBER OF SEQ ID NOS: 911

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 344

; LENGTH: 171

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-770-696-344

Query Match 73.3%; Score 15.4; DB 10; Length 171;

Best Local Similarity 94.1%; Pred. No. 1.7e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 AAGTTCTTGAGGCTCC 21

||||||| ||||| |||||

Db 34 AAGTTCTTGAGGCTCC 50

RESULT 12

US-10-051-307-4

; Sequence 4, Application US/10051307

; Patent No. US20020170095A1

; GENERAL INFORMATION:

; APPLICANT: DAI, ZIYU

; APPLICANT: SHI, LIFANG

; APPLICANT: HOOKER, BRIAN S.

; TITLE OF INVENTION: GENE PROMOTERS ISOLATED FROM POTATO AND USE THEREOF

; FILE REFERENCE: 059440/0141

; CURRENT APPLICATION NUMBER: US/10/051,307

; CURRENT FILING DATE: 2002-01-22

; PRIOR APPLICATION NUMBER: 60/263,224

; PRIOR FILING DATE: 2001-01-23

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 1175

; TYPE: DNA

; ORGANISM: Solanum tuberosum

US-10-051-307-4

Query Match 73.3%; Score 15.4; DB 9; Length 1175;

Best Local Similarity 94.1%; Pred. No. 2e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAAGTTTCTTGAGGCTC 20  
Db 139 GAAGTTTCTTGAGGCTC 155

## RESULT 13

US-09-938-842A-48  
; Sequence 48, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SAME, AND METHODS OF USE  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR APPLICATION NUMBER: 2001-08-24  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 48  
; LENGTH: 2979  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-48

Query Match 73.3%; Score 15.4; DB 9; Length 2979;  
Best Local Similarity 94.1%; Pred. No. 2.1e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAAGTTTCTTGAGGCTC 20  
Db 584 GAAGTTTCTTGAGTCTC 600

## RESULT 14

US-09-960-253-175/c  
; Sequence 175, Application US/09960253  
; Patent No. US20020123619A1  
; GENERAL INFORMATION:  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Lodes, Michael J.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.556  
; CURRENT APPLICATION NUMBER: US/09/960,253  
; CURRENT FILING DATE: 2001-09-20  
; NUMBER OF SEQ ID NOS: 187  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 175  
; LENGTH: 7497  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-960-253-175

Query Match 73.3%; Score 15.4; DB 10; Length 7497;  
Best Local Similarity 94.1%; Pred. No. 2.3e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAAGTTTCTTGAGGCTC 20  
Db 5201 GAAGTTTCTTGAGGCTC 5185

## RESULT 15

US-10-044-090-359/c

; Sequence 359, Application US/10044090  
; Patent No. US20020137081A1  
; GENERAL INFORMATION:  
; APPLICANT: Olga Bandman  
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION  
; FILE REFERENCE: PA-0028 US  
; CURRENT APPLICATION NUMBER: US/10/044,090  
; CURRENT FILING DATE: 2002-01-09  
; NUMBER OF SEQ ID NOS: 850  
; SOFTWARE: PERL Program  
; SEQ ID NO 359  
; LENGTH: 7792  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. US20020137081A1 2171870CB1  
US-10-044-090-359

Query Match 73.3%; Score 15.4; DB 12; Length 7792;  
Best Local Similarity 94.1%; Pred. No. 2.3e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAAGTTTCTTGAGGCTC 20  
Db 5170 GAAGTTTCTTGAGGCTC 5154

Search completed: July 9, 2003, 02:22:33  
Job time: 122.338 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: July 8, 2003, 00:47:28 ; Search time 1064.65 Seconds  
(without alignments)  
319.453 Million cell updates/sec

Title: US-09-647-780A-18  
Perfect score: 21  
Sequence: 1 ccggaagtcttcttgaggctcc 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

- 1: em\_estba:\*\*
- 2: em\_esthum:\*\*
- 3: em\_estin:\*\*
- 4: em\_estmd:\*\*
- 5: em\_estov:\*\*
- 6: em\_estpl:\*\*
- 7: em\_estro:\*\*
- 8: em\_htc:\*\*
- 9: gb\_est1:\*\*
- 10: gb\_est2:\*\*
- 11: gb\_htc:\*\*
- 12: gb\_est3:\*\*
- 13: gb\_est4:\*\*
- 14: gb\_est5:\*\*
- 15: em\_estfun:\*\*
- 16: em\_estom:\*\*
- 17: gb\_gss:\*\*
- 18: em\_gss\_hum:\*\*
- 19: em\_gss\_inv:\*\*
- 20: em\_gss\_pln:\*\*
- 21: em\_gss\_vrt:\*\*
- 22: em\_gss\_fun:\*\*
- 23: em\_gss\_mam:\*\*
- 24: em\_gss\_mus:\*\*
- 25: em\_gss\_other:\*\*
- 26: em\_gss\_pro:\*\*
- 27: em\_gss\_rtd:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	17.8	84.8	416	10 AW845925	AW845925 QV4-CT014
C 2	17.8	84.8	463	14 B0636494	B0636494 hd10d06.y
C 3	17.4	82.9	732	17 BH541990	BH541990 B06XG49TR
C 4	17	81.0	741	17 BH140462	BH140462 ZMMBB000
C 5	16.8	80.0	250	9 AI071748	AI071748 UT-R-C2-n
C 6	16.8	80.0	250	10 BB230340	BB230340 BB230340

C 7	16.8	80.0	299	10 BB449612	BB449612 BB449612
C 8	16.8	80.0	305	10 BB548769	BB548769 BB548769
C 9	16.8	80.0	471	17 AZ484779	AZ484779 LM0311C02
C 10	16.8	80.0	548	14 BQ994599	BQ994599 QG7K21.y
C 11	16.8	80.0	567	12 BF070228	BF070228 st08f11.y
C 12	16.8	80.0	579	9 AA521007	AA521007 aa70e02.s
C 13	16.8	80.0	645	13 BI065269	BI065269 pgt1n.pk0
C 14	16.8	80.0	684	12 BG351521	BG351521 109E03 Ma
C 15	16.8	80.0	705	10 AW148283	AW148283 dal4h03.x
C 16	16.8	80.0	740	14 BQ988381	BQ988381 QGFL4M09
C 17	16.8	80.0	756	14 BQ988193	BQ988193 QGFL4E06
C 18	16.8	80.0	802	17 BH071641	BH071641 RPCI-24-3
C 19	16.8	80.0	954	14 BQ938634	BQ938634 AGENCOURT
C 20	16.4	78.1	284	9 AA960891	AA960891 on97h10.s
C 21	16.4	78.1	356	13 BI478067	BI478067 dal09h01
C 22	16.4	78.1	396	9 AA570749	AA570749 nf21a08.s
C 23	16.4	78.1	432	14 T96412	T96412 yea3e06.s1
C 24	16.4	78.1	497	17 AQ828497	AQ828497 HS-5255-B
C 25	16.4	78.1	503	17 AQ060080	AQ060080 CIT-HSP-2
C 26	16.4	78.1	555	9 AI651388	AI651388 wb05f07.x
C 27	16.4	78.1	559	11 AK015246	AK015246 Mus muscu
C 28	16.4	78.1	626	10 AV713343	AV713343 AV713343
C 29	16.4	78.1	630	10 BB463960	BB463960 BB463960
C 30	16.4	78.1	633	9 AA538091	AA538091 vj56803.r
C 31	16.4	78.1	656	13 BJ091864	BJ091864 BJ091864
C 32	16.4	78.1	953	12 BG687526	BG687526 602639318
C 33	16.2	77.1	186	10 AW800214	AW800214 MR2-UM006
C 34	16.2	77.1	198	17 AZ226274	AZ226274 RPCI-23-9
C 35	16.2	77.1	212	14 BQ332470	BQ332470 PM3-ET026
C 36	16.2	77.1	212	14 BQ332491	BQ332491 PM3-ET026
C 37	16.2	77.1	238	9 AI858370	AI858370 wl37c06.x
C 38	16.2	77.1	324	9 AV116750	AV116750 AV116750
C 39	16.2	77.1	333	10 AW418979	AW418979 hal6c08.x
C 40	16.2	77.1	379	9 AI949943	AI949943 wq04e12.x
C 41	16.2	77.1	441	9 AA040036	AA040036 zk47f02.s
C 42	16.2	77.1	469	12 BG131509	BG131509 EST464401
C 43	16.2	77.1	480	12 BG143603	BG143603 ut57g12.x
C 44	16.2	77.1	498	12 BG294354	BG294354 602391451
C 45	16.2	77.1	498	13 BJI97096	BJI97096 BJI97096

ALIGNMENTS

RESULT 1	AW845925/C	416 bp	mRNA	linear	EST 19-MAY-2000
LOCUS	QV4-CT0141-170999-003-e05 CT0141	Homo sapiens cDNA, mRNA sequence.			
DEFINITION	AW845925				
ACCESSION	AW845925.1	GI:7941442			
VERSION	EST				
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 416)				
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.				
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)				
MEDLINE	20202663				
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001				

Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tbl=st2-QV4-CT0141-170  
 999-003-e03st3-1999-09-17&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 45  
 High quality sequence stop: 302.

## FEATURES

Location/Qualifiers  
 1..416  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="CT0141"  
 /dev\_stage="Adult"

/note="Organ: colon; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 104 a 107 c 132 g 73 t  
 ORIGIN

Query Match 84.8%; Score 17.8; DB 10; Length 416;  
 Best Local Similarity 90.5%; Pred. No. 3.2e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGAAGTTTCTTGAGGCTCC 21  
 ||||| ||| ||||| |||||  
 Db 300 CCGAAGTTCTCTGAGGCTCC 280

## RESULT 2

BQ636494/c

## LOCUS

DEFINITION BQ636494 463 bp mRNA EST 15-JUL-2002  
 hd10d06.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he

ACCESSION BQ636494

VERSION BQ636494.1 GI:21760953

KEYWORDS EST.

SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 463)

Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A., Touchman

,J.W., Bouffard,G., Smith,D. and Peterson,K.

Expressed sequence tag analysis of human retina for the NEIBank

Project: Retbindin, an abundant, novel retinal cDNA and alternative

splicing of other retina-preferred gene transcripts

Mol..Vis. 8 (4), (2002) In press

Contact: Wistow G

Section on Molecular Structure and Function

National Eye Institute

6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 10 row: d column: 06

Seq primer: M13Rpl reverse primer (AB1).

Location/Qualifiers

## FEATURES

1..463  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="hd10d06"

/note="Human Retina cDNA (Un-normalized, unamplified  
 ); hd/he"

/tissue\_type="Retina"

/dev\_stage="Adult"

/lab\_host="EMDH10B"

/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue  
 was dissected from two 80 year old donors with no observed

eye disease. 100ug of total RNA was used for library construction. A directionally cloned cDNA library in the pSPORT1 vector (Life Technologies) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the Superscript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a NOT I primer-adaptor [5'-pgactagttctagatcgagcgccgcc(T)15-3' ]. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

BASE COUNT 110 a 113 c 150 g 90 t  
 ORIGIN

Query Match 84.8%; Score 17.8; DB 14; Length 463;

Best Local Similarity 90.5%; Pred. No. 3.4e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGAAGTTTCTTGAGGCTCC 21

||||| ||| ||||| |||||

Db 305 CCGAAGTTCTTGAGGCTCC 285

## RESULT 3

BH541990

## LOCUS

DEFINITION BQXG49TR BQX Brassica oleracea genomic clone BQXG49, DNA  
 sequence.

ACCESSION BH541990

VERSION BH541990.1 GI:17793027

KEYWORDS GSS.

SOURCE Brassica oleracea.

## ORGANISM

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 732)

TOWN,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished (2001)

Other\_GSSs: BQXG49TF

Contact: Chris Town

## TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..732

/organism="Brassica oleracea"

/strain="TO1000DH3"

/db\_xref="taxon:3712"

/clone="BQXG49"

/clone\_lib="BQX"

/note="Vector: PHOS1; Site\_1: BstXI; 2-3 kb sheared  
 genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 212 a 153 c 173 g 194 t

## ORIGIN

Query Match 82.9%; Score 17.4; DB 17; Length 732;

Best Local Similarity 94.7%; Pred. No. 6.1e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCGAAGTTTCTTGAGGCT 19

||||| ||| ||||| |||||

Db 438 CCGAAGTTCTTGAGACT 456

## RESULT 4

BH140462/c

LOCUS BH140462 741 bp DNA linear GSS 07-AUG-2001  
DEFINITION ZMMBB0002F02r Maize B73 Zea mays genomic clone ZMMBB0002F02r, DNA sequence.  
ACCESSION BH140462  
VERSION BH140462.1 GI:15099523  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 741)  
Tomkins,J.P., Main,D., Goicoechea,J.L., Frisch,D.A. and Wing,R.A.  
TITLE A Deep-Coverage BAC Library for Maize  
JOURNAL Unpublished (2001)  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Class: BAC ends  
High quality sequence stop: 619.  
Location/Qualifiers  
1. 741  
/organism="Zea mays"  
/strain="B73"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBB0002F02r"  
/clone\_lib="Maize B73"  
/tissue\_type="Young leaves"  
/lab\_host="E. coli"  
/note="Vector: pCUGIBAC-1; Site\_1: HindIII; Site\_2: NotI; For more details on library preparation, ordering clones and sequence analysis see  
http://www.genome.clemson.edu/projects/stc/maize/ZMMBBb "

BASE COUNT 250 a 142 c 159 g 190 t  
ORIGIN  
Query Match 81.0%; Score 17; DB 17; Length 741;  
Best Local Similarity 100.0%; Pred.No. 9.6e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCCGAGTTCTTGAGG 17  
|||||  
Db 677 CCCGAGTTCTTGAGG 661  
RESULT 5  
AI071748 250 bp mRNA linear EST 05-JUL-1999  
LOCUS AI071748  
DEFINITION UI-R-C2-nj-b-10-0-UI.s1 UI-R-C2 Rattus norvegicus cDNA clone  
UI-R-C2-nj-b-10-0-UI 3', mRNA sequence.  
ACCESSION AI071748  
VERSION AI071748.1 GI:3397963  
KEYWORDS EST.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
1 (bases 1 to 250)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
COMMENT Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msorensen@blue.weeg.uiowa.edu  
The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized adult 8-day-Embryo library. cDNA Library Preparation: M. Fatima Bonaldo, Ph.D. Clone distribution: clones will be available through Research Genetics This clone is also available through the I.M.A.G.E. Consortium at LLNL (info@image.llnl.gov). IMAGE ID-1776381  
Seq primer: M13 Forward  
POLYA-No.  
Location/Qualifiers  
1. 250  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-C2-nj-b-10-0-UI"  
/clone\_lib="UI-R-C2"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; The UI-R-C2 library is a subtracted library derived from the UI-R-C1 library, which is a subtracted library derived from the UI-R-C0 library. The UI-R-C0 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C2) was constructed as follows: PCR amplified cDNA inserts from UI-R-C1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C2 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)."  
BASE COUNT 54 a 49 g 92 t  
ORIGIN  
Query Match 80.0%; Score 16.8; DB 9; Length 250;  
Best Local Similarity 90.0%; Pred.No. 8.5e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CCCGAGTTCTTGAGGCTC 20  
|||||  
Db 186 CCCGAGTTCTTGAGGTC 205  
RESULT 6  
BB230340 250 bp mRNA linear EST 03-JUL-2000  
LOCUS BB230340  
DEFINITION RIKEN full-length enriched, 3 days neonate thymus Mus musculus cDNA clone A630026D12.3', mRNA sequence.  
ACCESSION BB230340  
VERSION BB230340.1 GI:8900985  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 250)  
Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Horii,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,

RESULT 7	BB449612	299 bp	mrna	linear	EST 21-JUL-2000				
LOCUS	BB449612	RIKEN full-length enriched, 9 days embryo Mus musculus							
DEFINITION	cDNA clone D030072M21 3', mRNA sequence.								
ACCESSION	BB449612								
VERSION	BB449612.1	GI:9310647							
KEYWORDS	EST.								
SOURCE	house mouse.								
ORGANISM	Mus musculus								
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;								
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.								
	1 (bases 1 to 299)								
	Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci								
	,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,								
	Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,								
	Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,								
	Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.								
	, Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.								
	, Ono,T., Owa,C., Sato,H., Sakai,C., Sato,K., Shibata,K., Shibata								
	,Y., Shigemoto,Y., Shingagawa,A., Shiraki,T., Soabe,Y., Sugahara,Y.								
	, Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya								
	,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.								
	, Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino								
	,M., Muramatsu,M. and Hayashizaki,Y.								
TITLE	RIKEN Mouse ESTs (Konno,H., et al.)								
JOURNAL	Unpublished (2000)								
COMMENT	Contact: Yoshihide.Hayashizaki								
	Laboratory for Genome Exploration Research Group, RIKEN Genomic								
	Sciences Center(GSC), Yokohama Institute								
	The Institute of Physical and Chemical Research (RIKEN)								
	1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan								
	Tel: 81-45-503-9222								
	Fax: 81-45-503-9216								

Fax: 01-43-503-5249  
 Email: genome-resescg.riken.go.jp;  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Hayashizaki, Y., Okazaki, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

Location/Qualifiers  
 1. .299  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="D030072M21"  
 /clone\_lib="RIKEN full-length enriched, 9 days embryo"  
 /dev\_stage="9 days embryo"  
 /lab\_host="DH10B"  
 /notes="Site.1: Salt; Site.2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 5'  
 GAGAGAGAGCGCGCCGACCTCTTTTCTTTTCTT 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15'

FEATURES  
 source

GAGAGAGAGATTCGAGTTAATTTAAAAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified

```

BASE COUNT      64 a      74 c      48 g      113 t
ORIGIN
/dev_stage="2 days pregnant adult"
/lab_host="DH10B"
/note="Site_1: Sali; Site_2: BamHI; CDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGCGCGCCAACTCGAGTTAATTAATCCCGCCCCCCCC 3']. CDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGATTCGAGTTAATTAATTAATCCCGCCCCCCCC 3']. CDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT      85 a      81 c      56 g      83 t
ORIGIN
Query Match      80.0%; Score 16.8; DB 10; Length 305;
Best Local Similarity 90.0%; Pred. No. 9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCGAAGTTCTTCGAGGCTC 20
    ||||| ||||| ||||| ||
Db 142 CCCGAAGTTACTTGAGGPTC 123

RESULT 9
A2484779/c 471 bp DNA linear GSS 05-OCT-2000
LOCUS IM0311C02R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0311C02 R, DNA sequence.
ACCESSION A2484779
VERSION A2484779.1 GI:10649953
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 471)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mamoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112 USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0311 row: C column: 02
Seq primer: CACACAGGAACAGCATATCACC
Class: plasmid ends
High quality sequence stop: 471.
Location/Qualifiers
1. 471
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0311C02"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson

```

```

BASE COUNT      64 a      74 c      48 g      113 t
ORIGIN
/dev_stage="2 days pregnant adult"
/lab_host="DH10B"
/note="Site_1: Sali; Site_2: BamHI; CDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGCGCGCCAACTCGAGTTAATTAATCCCGCCCCCCCC 3']. CDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGATTCGAGTTAATTAATTAATCCCGCCCCCCCC 3']. CDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT      85 a      81 c      56 g      83 t
ORIGIN
Query Match      80.0%; Score 16.8; DB 10; Length 299;
Best Local Similarity 90.0%; Pred. No. 9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCGAAGTTCTTCGAGGCTCC 21
    ||||| ||||| ||||| ||
Db 55 CCGAAGTTCTTCGAGGCTCC 36

RESULT 8
BB548769/c 305 bp mRNA linear EST 31-JUL-2000
LOCUS BB548769 RIKEN full-length enriched, 2 days pregnant adult female
DEFINITION oviduct Mus musculus cDNA clone E230012D10 3', mRNA sequence.
ACCESSION BB548769
VERSION BB548769.1 GI:9620197
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 305)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.,
Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
Ono,T., Owa,C., Sakai,C., Sato,K., Shibata,K., Shibata,Y.,
Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.,
Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya
,T., Tsunoda,Y., Watabiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.,
Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
,M., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Konno,H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermotabilization and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitzunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.
Location/Qualifiers
1. 305
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="E230012D10"
/clone_lib="RIKEN full-length enriched, 2 days pregnant
adult female oviduct"
/sex="female"
/tissue_type="oviduct"

```

Laboratory Mouse DNA Resource  
(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWB42 (gll14732114,gbAF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adaptored vector DNA and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

BASE COUNT 127 a 112 c 90 g 142 t

Query Match 80.0%; Score 16.8; DB 17; Length 471;  
Best Local Similarity 90.0%; Pred. No. 1e+03; Indels 0; Gaps 0;  
Matches 18; Conservative 0; Mismatches 2;

OY 2 CCGAAGTTCTTCGAGCTCC 21

1 |||||

Db 448 CAGAAGTTCTTCGAGCTCC 429

# RESULT 10

B0994599/c

LOCUS

DEFINITION

QGF7K21.yg.ab1 QG\_EFGHJ lettuce serriola Lactuca sativa cDNA clone

QGF7K21 mRNA sequence.

ACCESSION B0994599

VERSION B0994599.1

KEYWORDS

SOURCE

ORGANISM

Lactuca sativa

Lactuca sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;

Lactuca.

REFERENCE

1. (bases 1 to 548)

AUTHORS

Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,

Lin,H., van Damme,M., Lavelle,D., Chevallier,P., Ziegler,J., Ellison

,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,

Church,S., Jackson,L. and Bradford,K.

Lettuce and Sunflower ESTs from the Compositae Genome Project

<http://comgenomics.ucdavis.edu/>

Unpublished (2002)

Contact: Alexander Kozik [R.W.Michelmore]

Department of Vegetable Crops, R.W.Michelmore Lab

University of California at Davis (UCD)

Asmundson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]

belongs to contig QG\_CA\_Contig5566, see <http://cgdb.ucdavis.edu/>

for details.

Plate: QGF7 row: K column: 21

Location/Qualifiers

1. 548

/organism="Lactuca sativa"

/cultivar="L.serriola"

/db\_xref="taxon:4236"

/clone="QGF7K21"

/clone\_lib="QG\_EFGHJ lettuce serriola"

/lab\_host="E.coli"

/note="Vector: pBRCDNASFIAB; The library was constructed

from 10 different sources of RNA from a single genotype.

Separate cDNAs were generated using primers that

incorporated unique 5' and 3' tags to distinguish each

source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgdb.ucdavis.edu/>

TAG\_L1B-QG\_EFGHJ lettuce serriola

TAG\_TISSUE=flowers environmental stress

TAG\_SEQ=CGAATCGGG

BASE COUNT 170 a 88 c 148 g 141 t

ORIGIN

Query Match 80.0%; Score 16.8; DB 14; Length 548;

Best Local Similarity 90.0%; Pred. No. 1.1e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CCCGAAGTTCTTCGAGCTC 20

1 |||||

Db 399 CCCGAAGTTCTTCGAGATC 380

# RESULT 11

BFO70228

LOCUS

DEFINITION

st08f11.y1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:

Gm-cl065-742 5' similar to TR-Q92QA1 Q92QA1 F13K3.13 PROTEIN ;

mRNA sequence.

ACCESSION BFO70228

VERSION BFO70228.1

KEYWORDS

SOURCE

ORGANISM

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

REFERENCE

1. (bases 1 to 567)

AUTHORS

Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corvett,V., Khanna

,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk

,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann

,R., Waterston,R. and Wilson,R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact via email: [ccu@resgen.com](mailto:ccu@resgen.com)

High quality sequence stop: 406

Location/Qualifiers

1. 567

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-742"

/clone\_lib="Gm-cl065"

/tissue\_type="germinating shoots"

/lab\_host="DH10B"

/note="Vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2:

XhoI; The cDNA library was constructed from mRNA isolated

germinating shoots of the cultivar Williams. The seeds

were allowed to germinate for 24 hours prior to being

cold stressed for 2 days at 4C. Complementary DNA was

synthesized from mRNA using a primer consisting of a

poly(dT) sequence with a XhoI restriction site. EcoRI

adapters were ligated to the blunt-ended cDNA fragments

followed by XhoI digestion. The cDNA fragments were

directionally cloned into the EcoRI-XhoI restriction site



of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 147 a 149 c 111 g 159 t 1 others  
ORIGIN

Query Match 80.0%; Score 16.8; DB 12; Length 567;  
Best Local Similarity 90.0%; Pred. No. 1.1e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGGAAGTTCTTGGGCTC 20  
||||| ||||||| |||||  
DB 388 CCGGATGTTCTTGGGCTC 407

## RESULT 12

AA521007 579 bp mRNA linear EST 20-AUG-1997  
LOCUS aa70e02.sl NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:826298 3',  
DEFINITION mRNA sequence.

ACCESSION AA521007  
VERSION AA521007.1 GI:2261550  
KEYWORDS EST.  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 579)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

## JOURNAL

Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-re@mail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,  
Ph.D., Gerald Marti, M.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldi, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/Dbrrp/image.html](http://www-bio.llnl.gov/Dbrrp/image.html)

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 146.

Location/Qualifiers

1..579

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="NCI\_CGAP\_GCB1"

/tissue\_type="germinal center B cell"

/lab\_host="DH10B"

/note="Vector: p773D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was prepared from human tonsillar cells enriched for  
germinal center B cells by flow sorting (CD20+, IgD+),  
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
primed with a Not I - oligo(GT) primer  
[5'-GTCTACCAATCTGAAGTGGAGCGGCTCATTTTCTTTT-3',  
J. Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified p773 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldi."

BASE COUNT 124 a 139 c 119 g 197 t  
ORIGIN

Query Match 80.0%; Score 16.8; DB 9; Length 579;  
Best Local Similarity 90.0%; Pred. No. 1.1e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGGAAGTTCTTGGGCTC 20  
||||| ||||||| |||||  
DB 388 CCGGATGTTCTTGGGCTC 407

## RESULT 13

BI065269/c

LOCUS

DEFINITION

cdna clone pgfln.pk004.b11 5', similar to no significant hits

(pLog(P) 4) G, mRNA sequence.

ACCESSION BI065269

VERSION BI065269.1 GI:14472791

KEYWORDS EST.

SOURCE

ORGANISM

Chicken.

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 645)

AUTHORS Cogburn, L.A., Morgan, R.W. and Burnside, J.

TITLE Chicken ESTs from fat

JOURNAL Unpublished (2001)

COMMENT Contact: Larry A. Cogburn

University of Delaware

Townsend Hall, Newark, DE 19717, USA

Fax: 302-831-1335

TEL: 302-831-2822

Email: cogburnudel.edu, www.chickest.udel.edu.

Location/Qualifiers

1..645

/organism="Gallus gallus"

/db\_xref="taxon:9031"

/clone\_lib="pgfln.pk004.b11"

/clone\_lib="normalized chicken fat cDNA library"

/sex="Male and female"

/tissue\_type="fat"

/lab\_host="E.coli EMDH10B"

/note="Vector: pSPORT1"

BASE COUNT 142 a 161 c 175 g 150 t 17 others  
ORIGIN

Query Match 80.0%; Score 16.8; DB 13; Length 645;  
Best Local Similarity 90.0%; Pred. No. 1.2e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGGAAGTTCTTGGGCTC 20  
||||| ||||||| |||||  
DB 405 CCGGATGTTCTTGGGCTC 386

## RESULT 14

BG351521/c

LOCUS

DEFINITION

109E03 Mature tuber. lambda ZAP Solanum tuberosum cDNA, mRNA

sequence.

ACCESSION BG351521

VERSION BG351521.1 GI:13180263

KEYWORDS EST.

SOURCE

ORGANISM

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 684)

AUTHORS Nielsen, K.L., Crookshanks, M., Emmersen, J. and Welinder, K.G.

TITLE EST-sequencing of mature potato tuber (Var. Kuras)

JOURNAL Unpublished (2000)

COMMENT Contact: Karen G. Welinder

Institut for bioteknologi

Aalborg Universitet

Sohngaardsholmsvej 49, 9000 Aalborg, Denmark

Tel: +45 96358467  
Fax: +45 98141808

Email: kgw@bio.auc.dk

Sequenced from the 5' end.  
High quality sequence stop: 684

POLYA-No.

#### FEATURES

##### Source

1. 684  
/organism="Solanum tuberosum"  
/cultivar="Field grown Kurat"  
/db\_xref="taxon:4113"  
/clone\_lib="Mature tuber lambda ZAP"  
/tissue\_type="Tuber"  
/note="Vector: Lambda ZAP"

191 a 129 c 167 g 197 t

#### BASE COUNT

##### ORIGIN

Query Match 80.0%; Score 16.8; DB 12; Length 684;

Best Local Similarity 90.0%; Pred. No. 1.2e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCGAAGTTTCTTGAGGCTC 20

|||||

Db 227 CCCGAAGTTTCTTGAGATTC 208

#### RESULT 15

##### LOCUS

##### DEFINITION

AW148283 705 bp mRNA linear EST 22-JUN-2000  
da14h03.x1 normalized Xenopus laevis gastrula laevis cDNA  
clone XENOPUS\_SOURCE:ID:xlnga001h06 3' similar to SW:TPR\_HUMAN  
P12270 NUCLEOPROTEIN TPR. [1]; mRNA sequence.

##### ACCESSION

##### VERSION

##### KEYWORDS

##### SOURCE

##### ORGANISM

Xenopus laevis  
African clawed frog.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;

##### REFERENCE

##### AUTHORS

1 (bases 1 to 705)  
Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D.,  
Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person,  
B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,  
Waterston, R. and Wilson, R.  
WashU Xenopus EST project, 1999  
Unpublished (1999)

##### TITLE

##### JOURNAL

##### COMMENT

Contact: Sandy Clifton, Ph.D.  
WashU Xenopus EST project, 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@wustl.edu

Library constructed by Bruce Blumberg

Library normalized by Jihwan Song

Clone Sequencing by: Washington University Genome Sequencing Center

DNA distribution: Xenopus clone distribution information for

this library can be found through Research Genetics, visit their

web page at: <http://www.resgen.com/>

Seq primer: -400P from Gibco

High quality sequence stop: 476.

Location/Qualifiers

#### FEATURES

##### source

1. 705  
/organism="Xenopus laevis"  
/db\_xref="taxon:8355"  
/clone="XENOPUS\_SOURCE:ID:xlnga001h06"  
/clone\_lib="normalized Xenopus laevis gastrula"  
/tissue\_type="gastrula (stages 10.5, 11.5 mixed)"  
/lab\_host="Top-10 F"  
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
XhoI; cDNA was prepared from 2ug of poly A+ RNA (equal  
parts from stage 10.5 and stage 11.5 gastrulae).

EcoRI-XhoI cut cDNA was then ligated into Unizap-XR  
(Stratagene) with EcoRI at the 5' end and XhoI at the 3'  
end. SS-library phagemids were prepared by mass excision  
from the original library and normalized by hybridization  
to biotinylated driver (prepared from the same library by  
PCR) to Cot-omega of 11. After removal of hybrids and  
excess driver by streptavidin sepharose chromatography,  
the ss-phagemids were made double stranded and  
electroporated into Top-10 F'. Original library  
constructed by Bruce Blumberg (Cho et al. 1991 Cell 67,  
1111-1120). Normalized by Jihwan Song (Song, Cho and  
Blumberg, unpublished). Note: This is a Xenopus Gene  
Collection (XGC) library."

BASE COUNT 148 a 139 c 192 g 223 t 3 others

#### ORIGIN

Query Match 80.0%; Score 16.8; DB 10; Length 705;

Best Local Similarity 90.0%; Pred. No. 1.2e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCGAAGTTTCTTGAGGCTC 20

|||||

Db 536 CCTGTAGTTTCTTGAGGCTC 555

Search completed: July 8, 2003, 09:22:09

Job time : 1069.8 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 00:45:53 ; Search time 209.093 Seconds  
(without alignments).  
2644.537 Million cell updates/sec

Title: US-09-647-780A-19  
Perfect score: 19  
Sequence: 1 gatcgctacctgactac 19

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg\_hum.\*
- 31: em.htg\_inv.\*
- 32: em.htg\_other.\*
- 33: em.htg\_mus.\*
- 34: em.htg\_pln.\*
- 35: em.htg\_rod.\*
- 36: em.htg\_man.\*
- 37: em.htg\_vrt.\*
- 38: em.sy.\*
- 39: em.htgo\_hum.\*
- 40: em.htgo\_mus.\*
- 41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	19	100.0	19	6	AX014719	Sequence
2	19	100.0	2765	6	AX014701	Sequence
3	19	100.0	174953	2	AC094732	Rattus no
4	17.4	91.6	2583	10	AF302075	Mus muscu
5	17.4	91.6	2601	10	AF157106	Mus muscu
6	17.4	91.6	2652	10	AF302076	Mus muscu
7	17.4	91.6	2694	10	AF302077	Mus muscu
8	17.4	91.6	2892	10	AF157105	Mus muscu
9	17.4	91.6	2925	6	AX033272	Sequence
10	17.4	91.6	2925	10	AF176569	Mus muscu
11	17.4	91.6	208249	2	AL607032	Mus muscu
12	17	89.5	233488	2	AC122228	Mus muscu
13	16.4	86.3	193363	2	AC090432	Mus muscu
14	16.4	86.3	204867	2	AC121567	Mus muscu
15	16.4	86.3	206132	2	AC115760	Mus muscu
16	16.4	86.3	318372	2	AC124524	Mus muscu
17	15.8	83.2	2076	6	AX146976	Sequence
18	15.8	83.2	2232	6	AX139864	Sequence
19	15.8	83.2	2262	6	AX146980	Sequence
20	15.8	83.2	2340	6	AX146978	Sequence
21	15.8	83.2	2340	6	AX473102	Sequence
22	15.8	83.2	2636	6	AX139743	Sequence
23	15.8	83.2	2663	6	AX139745	Sequence
24	15.8	83.2	2676	6	AX033274	Sequence
25	15.8	83.2	2714	6	AX139741	Sequence
26	15.8	83.2	2784	9	AF336981	Homo sapi
27	15.8	83.2	2850	9	AK093058	Homo sapi
28	15.8	83.2	2893	6	AX356951	Sequence
29	15.8	83.2	2893	6	AX463057	Sequence
30	15.8	83.2	2953	6	AX473100	Sequence
31	15.8	83.2	2975	6	AX356955	Sequence
32	15.8	83.2	2975	6	AX463058	Sequence
33	15.8	83.2	34521	9	AL731535	Human DNA
34	15.8	83.2	60144	2	AC019798	Drosophill
35	15.8	83.2	65229	8	AP003705	Oryza sat
36	15.8	83.2	91685	2	AC005121	Drosophill
37	15.8	83.2	105227	2	AP005296	Oryza sat
38	15.8	83.2	105713	2	AL390200	Homo sapi
39	15.8	83.2	108795	9	AC114483	Homo sapi
40	15.8	83.2	110000	2	AC055726_1	Continuation (2 of
41	15.8	83.2	137767	2	AC093094	Oryza sat
42	15.8	83.2	141905	2	AC110101	Rattus no
43	15.8	83.2	148041	2	AC021320	Homo sapi
44	15.8	83.2	149291	10	AC116328	Mus muscu
45	15.8	83.2	154736	9	AL139246	Human DNA

ALIGNMENTS

RESULT 1  
AX014719  
LOCUS AX014719  
DEFINITION Sequence 19 from Patent WO953077.  
ACCESSION AX014719  
VERSION AX014719.1 GI:10040992  
KEYWORDS synthetic construct,  
artificial construct,  
artificial sequences.  
ORGANISM 1 (bases 1 to 19)  
REFERENCE Schwartz,J.C., Gros,C., Ouimet,T., Rose,C., Bonhomme,M.C. and  
AUTHORS Facchinetti,P.  
TITLE Novel nep ii membrane metalloprotease and its use for screening  
inhibitors useful in therapy

linear PAT 07-SEP-2000

```

JOURNAL Patent: WO 9953077-A 19 21-OCT-1999;
INST NAT SANTE RECH MED (FR); SCHWARTZ JEAN CHARLES (FR); GROS
CLAUDE (FR); OUMET TANJA (FR); ROSE CHRISTIANE (FR); BONHOMME
MARIE CHANTAL (FR); FACCHINETTI PATRICIA (FR)

FEATURES
Source
Location/Qualifiers
1..19
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"
4 a 7 c 4 g 4 t

BASE COUNT
ORIGIN
100.0%; Score 19; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCGGCTACCCCTGACTAC 19
|||||
Db 1 GATCGGCTACCCCTGACTAC 19

RESULT 2
AX014701 2765 bp DNA linear PAT 07-SEP-2000
LOCUS
DEFINITION
Sequence 1 from Patent WO9953077.
ACCESSION AX014701
VERSION
KEYWORDS
SOURCE
ORGANISM
black rat.
Rattus rattus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 2765)
Schwartz,J.C., Gros,C., Oumet,T., Rose,C., Bonhomme,M.C. and
Facchinetti,P.
Novel nep ii membrane metalloprotease and its use for screening
inhibitors useful in therapy
Patent: WO 9953077-A 1 21-OCT-1999;
INST NAT SANTE RECH MED (FR); SCHWARTZ JEAN CHARLES (FR); GROS
CLAUDE (FR); OUMET TANJA (FR); ROSE CHRISTIANE (FR); BONHOMME
MARIE CHANTAL (FR); FACCHINETTI PATRICIA (FR)

FEATURES
Source
Location/Qualifiers
1..2765
/organism="Rattus rattus"
/db_xref="taxon:10117"
107..2431
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07576.1"
/db_xref="GI:10040976"
/translation="MGKSSVGMERADNCRRRLGFEVCGLLVLLTLLMGAIVTL
GYFISGKQLPLNSLLHVSHERTVKRVLRDSSQSDICTPPSCVIAAARLIQND
QSKPCDNPQYQAGSWLRRHVPETNSVSVDFILRDEVLKGVLEDSSVQHRPA
VKAKTLVRSQACQSVLEKRDSEPLNLVDMIGGWPVMDKNETGWELERQAV
LSQNRRLVLDLFWNDQNSRHVIYIDQPTLGPMSREYFEKDSHRVREAYLOFM
TSVATMLRDLNPGTDLVQEMAEQVHLHLETLANATVPQEKRHDTALYHRMGLEE
LYAVRLVRLDRIGLSQVDFEYVQVLENEEYVYCIPIYLENEIIDVFPQATLQN
LQERFLGNFTLQNLVSSVQVLENEEYVYCIPIYLENEIIDVFPQATLQN
TSVATMLRDLNPGTDLVQEMAEQVHLHLETLANATVPQEKRHDTALYHRMGLEE
IKRAFSDKSKIVSELIERSVFVNDLNMWDESKKQAEKALNIREQIGYPDY
ILEDNRHLDEYSLSLTFSEDLFFENGLQNLKNAORSLSKLRKYDQNLWILGAAYV
NAFYSPNRNLVFPQIPFPFQKQALNPGFGIMVGHEITHTFDGNDGRNFKN
GNLMDWSNFSRHFQSQCMYQISNFWELADNONGVSTIGTENADNGVQQA
YKALVLAEGGRDRLPGLNLTAYQALFFINTAQVWCGYRPEFAIQSIKTVDHSPK
YRVLSQNLNPGFSEAFHCPGSPMHPMNRCLW"

BASE COUNT 684 a 735 c 787 g 559 t
ORIGIN
100.0%; Score 19; DB 6; Length 2765;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCGGCTACCCCTGACTAC 19
|||||
Db 1 GATCGGCTACCCCTGACTAC 1630

RESULT 3
AC094732 174953 bp DNA linear HTG 20-DEC-2001
LOCUS
DEFINITION
Rattus norvegicus clone CH230-516, *** SEQUENCING IN PROGRESS ***
56 unordered pieces.
AC094732
HTG: HTGS PHASE1.
Rattus norvegicus
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 174953)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Correll,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Joudah,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovat,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louleghed,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonaite,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 174953)
Worley,K.C.
Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15624568.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GBGF

```

Center clone name: CH230-516  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329First call to  
 findPhrapList  
 Consensus quality: 152255 bases at least Q40  
 Consensus quality: 158448 bases at least Q30  
 Consensus quality: 164461 bases at least Q20  
 Estimated insert size: 155965; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-fp estimation  
 Quality coverage: 2.2x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 consists of 56 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1 13782: contig of 13782 bp in length  
 \* 13783 13882: gap of unknown length  
 \* 13883 23287: contig of 9405 bp in length  
 \* 23288 23387: gap of unknown length  
 \* 23388 28081: contig of 4694 bp in length  
 \* 28082 28181: gap of unknown length  
 \* 28182 33807: contig of 5626 bp in length  
 \* 33808 33907: gap of unknown length  
 \* 33908 39271: contig of 5364 bp in length  
 \* 39272 39371: gap of unknown length  
 \* 39372 44270: contig of 4899 bp in length  
 \* 44271 44370: gap of unknown length  
 \* 44371 47723: contig of 3353 bp in length  
 \* 47724 47823: gap of unknown length  
 \* 47824 53427: contig of 5604 bp in length  
 \* 53428 53527: gap of unknown length  
 \* 53528 57303: contig of 3776 bp in length  
 \* 57304 57403: gap of unknown length  
 \* 57404 62018: contig of 4615 bp in length  
 \* 62019 62118: gap of unknown length  
 \* 62119 66776: contig of 4558 bp in length  
 \* 66777 66776: gap of unknown length  
 \* 66777 70201: contig of 3425 bp in length  
 \* 70202 70301: gap of unknown length  
 \* 70302 74082: contig of 3781 bp in length  
 \* 74083 74182: gap of unknown length  
 \* 74183 78306: contig of 4124 bp in length  
 \* 78307 78406: gap of unknown length  
 \* 78407 81222: contig of 2816 bp in length  
 \* 81223 81322: gap of unknown length  
 \* 81323 84552: contig of 3230 bp in length  
 \* 84553 84652: gap of unknown length  
 \* 84653 88839: contig of 4187 bp in length  
 \* 88840 88939: gap of unknown length  
 \* 88940 92256: contig of 3317 bp in length  
 \* 92257 92356: gap of unknown length  
 \* 92357 94995: contig of 2639 bp in length  
 \* 94996 95095: gap of unknown length  
 \* 95096 98465: contig of 3370 bp in length  
 \* 98466 98565: gap of unknown length  
 \* 98566 101632: contig of 3067 bp in length  
 \* 101633 101732: gap of unknown length  
 \* 101733 104939: contig of 3207 bp in length  
 \* 104940 105039: gap of unknown length  
 \* 105040 109248: contig of 4209 bp in length  
 \* 109249 109348: gap of unknown length  
 \* 109349 112435: contig of 3087 bp in length  
 \* 112436 112535: gap of unknown length  
 \* 112536 114677: contig of 2141 bp in length  
 \* 114677 114776: gap of unknown length  
 \* 114777 117567: contig of 2791 bp in length  
 \* 117568 117667: gap of unknown length

117668 119209: contig of 1542 bp in length  
 \* 119210 119309: gap of unknown length  
 \* 119310 121797: contig of 2488 bp in length  
 \* 121798 121897: gap of unknown length  
 \* 121898 124293: contig of 2396 bp in length  
 \* 124294 124393: gap of unknown length  
 \* 124394 126428: contig of 2035 bp in length  
 \* 126429 126528: gap of unknown length  
 \* 126529 129525: contig of 2997 bp in length  
 \* 129526 129625: gap of unknown length  
 \* 129626 132334: contig of 2709 bp in length  
 \* 132335 132434: gap of unknown length  
 \* 132435 135274: contig of 2840 bp in length  
 \* 135275 135374: gap of unknown length  
 \* 135375 138074: contig of 2700 bp in length  
 \* 138075 138174: gap of unknown length  
 \* 138175 139985: contig of 1811 bp in length  
 \* 139986 140085: gap of unknown length  
 \* 140086 142273: contig of 2188 bp in length  
 \* 142274 142373: gap of unknown length  
 \* 142374 143598: contig of 1225 bp in length  
 \* 143599 143698: gap of unknown length  
 \* 143699 145434: contig of 1736 bp in length  
 \* 145435 145534: gap of unknown length  
 \* 145535 146985: contig of 1451 bp in length  
 \* 146986 147085: gap of unknown length  
 \* 147086 148099: contig of 1014 bp in length  
 \* 148100 148199: gap of unknown length  
 \* 148200 150915: contig of 2716 bp in length  
 \* 150916 151015: gap of unknown length  
 \* 151016 152501: contig of 1486 bp in length  
 \* 152502 152601: gap of unknown length  
 \* 152602 154010: contig of 1409 bp in length  
 \* 154011 154110: gap of unknown length  
 \* 154111 155758: contig of 1648 bp in length  
 \* 155759 155858: gap of unknown length  
 \* 155859 157622: contig of 1764 bp in length  
 \* 157623 157722: gap of unknown length  
 \* 157723 159428: contig of 1706 bp in length  
 \* 159429 159528: gap of unknown length  
 \* 159529 161209: contig of 1681 bp in length  
 \* 161210 161309: gap of unknown length  
 \* 161310 163413: contig of 2104 bp in length  
 \* 163414 163513: gap of unknown length  
 \* 163514 164702: contig of 1189 bp in length  
 \* 164703 164802: gap of unknown length  
 \* 164803 165998: contig of 1196 bp in length  
 \* 165999 166098: gap of unknown length  
 \* 166099 167412: contig of 1314 bp in length  
 \* 167413 167512: gap of unknown length  
 \* 167513 169231: contig of 1719 bp in length  
 \* 169232 169331: gap of unknown length  
 \* 169333 170534: contig of 1203 bp in length  
 \* 170535 170634: gap of unknown length  
 \* 170635 172047: contig of 1413 bp in length  
 \* 172048 172147: gap of unknown length  
 \* 172148 173509: contig of 1362 bp in length  
 \* 173510 173610: gap of unknown length  
 \* 173610 174953: contig of 1344 bp in length.

## FEATURES

Location/Qualifiers

Query Match 100.0%; Score 19; DB 2; Length 174953;  
 Best Local Similarity 100.0%; Pred. NO. 13;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCGGCTACCTGACTAC 19

Db 64036 GATCGGCTACCTGACTAC 64054

## RESULT 4

AF302075  
 LOCUS AF302075 2583 bp mRNA linear ROD\_11-JUN-2001

```

DEFINITION Mus musculus neprilysin-like peptidase alpha mRNA, complete cds.
ACCESSION AF302075
VERSION AF302075.1 GI:10505359
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 2583)
AUTHORS Shirotsani,K., Tsubuki,S., Iwata,N., Takaki,Y., Harigaya,W.,
Maruyama,K., Kiryu-Seo,S., Kiyama,H., Iwata,H., Tomita,T.,
Iwatsubo,T. and Saido,T.C.
TITLE Neprilysin degrades both amyloid beta peptides 1-40 and 1-42 most
rapidly and efficiently among thiorphan- and
phosphoramidon-sensitive endopeptidases
J. Biol. Chem. 276 (24), 21895-21901 (2001)
JOURNAL 21293028
MEDLINE 11278416
PUBMED
REFERENCE 2 (bases 1 to 2583)
AUTHORS Shirotsani,K. and Saido,T.C.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-2000) Proteolytic Neuroscience Laboratory, Brain
Science Institute, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama
351-0198, Japan
FEATURES             Location/Qualifiers
     source            1..2583
                        /organism="Mus musculus"
                        /db_xref="taxon:10090"
     CDS                25...2253
                        /note="endopeptidase"
                        /codon_start=1
                        /product="neprilysin-like peptidase alpha"
                        /protein_id="AG18446.1"
                        /db_xref="GI:10505360"
                        /translation="MVERAGWCRRKSPGFVEYGLMVLILLGLAIVTLGVFYSIALRD
                        SLSKSDICTPSCVIAAARILENNQSNPCENFYACGGWLRHHVLPETNSRYSVF
                        DILRDEVLTKGVLEDSOHRPAVEKATLYRSCMNQSVIEKRDESEPLLSVLKMGV
                        GWPVADKNETMGLKWELEQLAVLNSQFNRRVLDLFIWDDONSRRVYIDOPT
                        LGMPREYFQEDNNHKKRAYLEFMTSVATMLRKQDNLSKESAMVREMAEVLLEET
                        HLANATVPEKRDHVTALYHRMDLMEQLERFGLKFNWTLFIQNVLSSVEVELFPDEE
                        VVYGYPILENLEDDISYSARTMQNYLVWRLVDRIGLSORFKEARVDYRKALYGT
                        TVEEVRWCESVYVNSMESAVGSLYIKRAFSDKSTVRELIEKIRSFVDNLDELN
                        WDEESKKAQAKAMNIREQIGPYDILEDNKKHDEEYSSLTFEEDLYFENGLOLNLK
                        NNAQSLKLRKEDQNLALIGAAVVNAFYSPNRNQLVFPAGILOPPFESKDQPSLN
                        FGGIGVNGHEITHGDDNGRNFKNMLDWNSF SARHFQOQSCMIYQYGNFSWE
                        LADQNQVNGESLIGENIADNGVRQAYKAYLRWLADGGKQDRPLGLNLTIAQLFFIN
                        AQVWCGSYRPEFAVQSIKTVDHVSPLKYRVLGSLQNLPGFSEAFHCPRGSPMHPMKRCR
                        IW"
BASE COUNT 665 a 667 c 736 g 515 t
ORIGIN
Query Match. 91.6%; Score 17.4; DB 10; Length 2583;
Best Local Similarity 94.7%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATCGCGTACCCCTGACTAC 19
   ||| ||||| ||||| |||||
DB 1434 GATTGGCTACCCCTGACTAC 1452

RESULT 5
AF157106 2601 bp mRNA linear ROD 25-NOV-1999
LOCUS Mus musculus soluble secreted endopeptidase delta mRNA,
DEFINITION alternatively spliced product, complete cds.
ACCESSION AF157106
VERSION AF157106.1 GI:6467400
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 2601)
AUTHORS Shirotsani,K., Tsubuki,S., Iwata,N., Takaki,Y., Harigaya,W.,
Maruyama,K., Kiryu-Seo,S., Kiyama,H., Iwata,H., Tomita,T.,
Iwatsubo,T. and Saido,T.C.
TITLE Neprilysin degrades both amyloid beta peptides 1-40 and 1-42 most
rapidly and efficiently among thiorphan- and
phosphoramidon-sensitive endopeptidases
J. Biol. Chem. 276 (24), 21895-21901 (2001)
JOURNAL 21293028
MEDLINE 11278416
PUBMED
REFERENCE 2 (bases 1 to 2601)
AUTHORS Shirotsani,K. and Saido,T.C.

```

```

AUTHORS Ikeda,K., Emoto,N., Raharjo,S.B., Murchantari,Y., Saiki,K.,
Yokoyama,M. and Matsuo,M.
TITLE Molecular identification and characterization of novel
membrane-bound metalloprotease, the soluble secreted form of which
hydrolyzes a variety of vasoactive peptides
J. Biol. Chem. 274 (45), 32469-32477 (1999)
JOURNAL 20011457
MEDLINE 10542292
PUBMED
REFERENCE 2 (bases 1 to 2601)
AUTHORS Ikeda,K., Emoto,N. and Matsuo,M.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-1999) International Center for Medical Research,
Kobe University School of Medicine, 7-5-1 Kusunoki, Chuo, Kobe
6500017, Japan
FEATURES             Location/Qualifiers
     source            1..2601
                        /organism="Mus musculus"
                        /db_xref="taxon:10090"
                        /db_xref="GI:6467401"
     CDS                59..2287
                        /note="SEP(delta); metalloprotease; alternatively spliced"
                        /codon_start=1
                        /product="soluble secreted endopeptidase delta"
                        /protein_id="AAF13153.1"
                        /db_xref="GI:6467401"
                        /translation="MVERAGWCRRKSPGFVEYGLMVLILLGLAIVTLGVFYSIALRD
                        SLSKSDICTPSCVIAAARILENNQSNPCENFYACGGWLRHHVLPETNSRYSVF
                        DILRDEVLTKGVLEDSOHRPAVEKATLYRSCMNQSVIEKRDESEPLLSVLKMGV
                        GWPVADKNETMGLKWELEQLAVLNSQFNRRVLDLFIWDDONSRRVYIDOPT
                        LGMPREYFQEDNNHKKRAYLEFMTSVATMLRKQDNLSKESAMVREMAEVLLEET
                        HLANATVPEKRDHVTALYHRMDLMEQLERFGLKFNWTLFIQNVLSSVEVELFPDEE
                        TVEEVRWCESVYVNSMESAVGSLYIKRAFSDKSTVRELIEKIRSFVDNLDELN
                        WDEESKKAQAKAMNIREQIGPYDILEDNKKHDEEYSSLTFEEDLYFENGLOLNLK
                        NNAQSLKLRKEDQNLALIGAAVVNAFYSPNRNQLVFPAGILOPPFESKDQPSLN
                        FGGIGVNGHEITHGDDNGRNFKNMLDWNSF SARHFQOQSCMIYQYGNFSWE
                        LADQNQVNGESLIGENIADNGVRQAYKAYLRWLADGGKQDRPLGLNLTIAQLFFIN
                        AQVWCGSYRPEFAVQSIKTVDHVSPLKYRVLGSLQNLPGFSEAFHCPRGSPMHPMKRCR
                        IW"
BASE COUNT 655 a 681 c 748 g 517 t
ORIGIN
Query Match. 91.6%; Score 17.4; DB 10; Length 2601;
Best Local Similarity 94.7%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATCGCGTACCCCTGACTAC 19
   ||| ||||| ||||| |||||
DB 1468 GATTGGCTACCCCTGACTAC 1486

RESULT 6
AF302076 2652 bp mRNA linear ROD 11-JUN-2001
LOCUS Mus musculus neprilysin-like peptidase beta mRNA, complete cds.
DEFINITION
ACCESSION AF302076
VERSION AF302076.1 GI:10505361
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 2652)
AUTHORS Shirotsani,K., Tsubuki,S., Iwata,N., Takaki,Y., Harigaya,W.,
Maruyama,K., Kiryu-Seo,S., Kiyama,H., Iwata,H., Tomita,T.,
Iwatsubo,T. and Saido,T.C.
TITLE Neprilysin degrades both amyloid beta peptides 1-40 and 1-42 most
rapidly and efficiently among thiorphan- and
phosphoramidon-sensitive endopeptidases
J. Biol. Chem. 276 (24), 21895-21901 (2001)
JOURNAL 21293028
MEDLINE 11278416
PUBMED
REFERENCE 2 (bases 1 to 2652)
AUTHORS Shirotsani,K. and Saido,T.C.

```

```

/codon_start=1
/product="neprilysin-like peptidase gamma"
/protein_id="AAG18448.1"
/db_xref="GI:10505364"
/translation="MYERAGWCRKKSPGFVEYGLMWLLLLLLGAIYVTLGVFYSIALRDL
SSLSKSDICTPSCVIAAARLENNQSRNCFYQACGMWRHVIHVPETNSRYSVF
DIURDEVLVKGWLEWESQHPRAVEKATLRSCNQSVIERKSEPLSVIKRWVG
GHWADKWNETWGLKWEJLAVLNQFNRRLVLDLFITWDDQNSRHVITDQPT
GLMPSREYTFQEDNNHKVRKAYLEFMTSVATMLKQDNLSKESAMREEMAEVLELET
HLANATVPOEKRDHVTALYHRMDLMELQERFGLKDRVSLCSPGCGPTHSDQAGLELG
NPPASDRLVGLKGENWTLFIQNVLSSEVELFPDEEVVYVIGIPILENLEIDISYSA
RPMQNVLYRWLVLDRLGSLISQREKARVDYRKALYCTTVEEVREKCVSYVNSMESA
VGLSYIKRAFSDKSTSRRELIEKIRSFVDNLDELNMWDEESKKLAKQEKAMNIREQI
GYAVTLDNNKHLDBEYFSGLTFEDLYFENGQNLKNAQSLKKLREKVDQNLITII
GPDVNFATYSPNRNQYLPQFPFSDKDPQSLNFGGIGMWIHEITPHGDDNRG
NFKNGKMLYLDNFGSARHFQOQSQCMYIYQGNFELADLNQVNGFSTLGENIADNG
GVQOAKLRLWLADNGDKQDRLPGLNLTVAOLFIINVAQVMCGSYRPEFAVQSIIKTDV
HSPLKRYVLGSLQNLPGFSEAFHCRGSPHMPKRCRIW"
BASE COUNT      586 a      700 c      766 g      542 t
ORIGIN
      1  GATCGGCTACCCGTGACTAC 19
      ||| ||||| ||||| |||||
Ddb 1545 GATTGGCTACCCGTGACTAC 1563

Query Match      91.6%; Score 17.4; DB 10; Length 2694;
Best Local Similarity 94.7%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 8
AF157105 LOCUS      2892 bp      mRNA      linear      ROD 25-NOV-1999
DEFINITION Mus musculus soluble secreted endopeptidase mRNA, complete cds.
ACCESSION AF157105
VERSION AF157105.1 GI:6467398
KEYWORDS Mus musculus
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2892)
AUTHORS Ikeda,K., Emoto,N., Raharjio,S.B., Nurhantari,Y., Saiki,K.,
Yokoyama,M. and Matsuo,M.
TITLE Molecular identification and characterization of novel
membrane-bound metalloprotease, the soluble secreted form of which
hydrolyzes a variety of vasoactive peptides
J. Biol. Chem. 274 (45), 32469-32477 (1999)
JOURNAL 200111457
MEDLINE 10542292
PUBMED
REFERENCE 2 (bases 1 to 2892)
AUTHORS Ikeda,K., Emoto,N. and Matsuo,M.
DIRECT SUBMISSION
TITLE Submitted (08-JUN-1999) International Center for Medical Research,
Kobe University School of Medicine, 7-5-1 Kusunoki, Chuo, Kobe
6500017, Japan
FEATURES
source Location/Qualifiers
      1..2892
      /organism="Mus musculus"
      /db_xref="taxon:10090"
      281..2578
      /note="SEP; metalloprotease"
      /codon_start=1
      /product="soluble secreted endopeptidase"
      /protein_id="AAG13152.1"
      /db_xref="GI:6467399"
      /translation="MYERAGWCRKKSPGFVEYGLMWLLLLLLGAIYVTLGVFYSIGKGL
PLTSLHFSWDERTVVYKRALRDSLSKSDICTPSCVIAAARLENNQSRNCFNCEFY
QYACGMWRHVIHVPETNSRYSVFDILRLDEVLVKGWLESTQHPRAVEKATLVRS
CMNQSVIERKDSPLSVYIAKMGVGPVALDKWNETMGLKWELEKQALAVLMSQSNRRVL
IDFIWDDQNSRHVITDQPTGLMPSREYTFQEDNNHKVRKAYLEFMTSVATMLKQDN
LNSKESAMREEMAEVLELETNFKNGKMLYLDNFGSARHFQOQSQCMYIYQGNFELADN
QVNGFSTLGENIADNGGVQOAKLRLWLADNGDKQDRLPGLNLTVAOLFIINVAQVMCG
SYRPEFAVQSIIKTDVFNWTLFIQNVLSSEVELFPDEEVVYVIGIPILENLEIDISY
SARTMONTLVWRLVRLV"
CDS

```



RIGLSQRFKEARVDYRKALYGTVEEVNRWCYVNSNMESAVGSLYIKRAFSDKS  
KSTVRELIEKIRSVFVDNLDELNMDESKKAKOAKMNIHQIGYDPYIILEDNNKHL  
DEEYSSLTFFEDLYFENGLONLKNNAQBSLKLREKVDONLWIGAAVVNAFYSPNRN  
QIVFPAGILOPPFFSKDQPSLNFSGGIGVIGHEITHGDFDNGRNFKNMGLDWSN  
FSARHFQOQSCMIYQGNFSEWELADNONVNGFSSLGENIADNGVROQAYKAYLRWLA  
DGGKDRPLGLNTYALQFFINAYQVWCGSYRPFVAVQSIKTDVHSPLKRYVLGSLQN  
LPGFSEAFHCPRGSPHMPMKRCRIW"

BASE COUNT 717 a 774 c 827 g 574 t  
ORIGIN

Query Match 91.6%; Score 17.4; DB 10; Length 2892;  
Best Local Similarity 94.7%; Pred. No. 1.5e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATCGGCTACCTGACTAC 19  
Db 1759 GATTGGCTACCTGACTAC 1777

RESULT 9  
AX033272  
LOCUS AX033272 2925 bp DNA linear PAT 21-SEP-2000  
DEFINITION Sequence 12 from Patent WO0047750.  
ACCESSION AX033272  
VERSION AX033272.1 GI:10280087  
KEYWORDS house mouse.  
SOURCE Mus musculus

ORGANISM  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS Boileau, G. and Desrosiers, L.  
TITLE New metalloproteases of the neprilysin family  
JOURNAL Patent: WO 0047750-A.12 17-AUG-2000;  
BOILEAU GUY (CA); DESGROSEILLERS LUC (CA); UNIVERSITE DE MONTREAL  
(CA)

FEATURES  
source 1..2925  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
332..2629  
CDS  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC09977.1"  
/db\_xref="GI:10280088"

/translation="MVRAGWCRKSPGVEYGLMVLVLLLLGAVTLGVFYSIGKQL  
PLTSLHFSWDETVVYKRALRDSLSKSDICTPSCVIAAARIENMDQSRNCFNY  
QYACGGWLRHVIPETNSRYSVFDILRDELEVILKGVLEDSTSOHRPAVERAKTYLRS  
CMQSVTEKRDSEPLSVLKMVGMPVAMDKWNETMGLKWELEQLAVLNSQFNRLV  
IDLFIWDDQNSRHVIYIDOPTLGMPSREYVFOEDNNHKVKAYLEFMTSVATMLRK  
DQNLKESAMVREMAEYVLETHLANATVPEKRHDVTALYHRMDLMELOERFGLKG  
FNWTLFTONVLSSEVEVLPDEEVVYVIGPYLENLEIDISYARTMQLVWRVLVD  
RIGLSQRFKEARVDYRKALYGTVEEVNRWCYVNSNMESAVGSLYIKRAFSDKS  
KSTVRELIEKIRSVFVDNLDELNMDESKKAKOAKMNIHQIGYDPYIILEDNNKHL  
DEEYSSLTFFEDLYFENGLONLKNNAQBSLKLREKVDONLWIGAAVVNAFYSPNRN  
QIVFPAGILOPPFFSKDQPSLNFSGGIGVIGHEITHGDFDNGRNFKNMGLDWSN  
FSARHFQOQSCMIYQGNFSEWELADNONVNGFSSLGENIADNGVROQAYKAYLRWLA  
DGGKDRPLGLNTYALQFFINAYQVWCGSYRPFVAVQSIKTDVHSPLKRYVLGSLQN  
LPGFSEAFHCPRGSPHMPMKRCRIW"

BASE COUNT 710 a 797 c 836 g 582 t  
ORIGIN

Query Match 91.6%; Score 17.4; DB 6; Length 2925;  
Best Local Similarity 94.7%; Pred. No. 1.5e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATCGGCTACCTGACTAC 19  
Db 1810 GATTGGCTACCTGACTAC 1828

RESULT 10  
AF176569  
LOCUS AF176569 2925 bp mRNA linear ROD 11-MAY-2000  
DEFINITION Mus musculus neprilysin-like metallopeptidase 1 (N11) mRNA,  
complete cds.  
ACCESSION AF176569  
VERSION AF176569.1 GI:7769082  
KEYWORDS  
SOURCE Mus musculus

ORGANISM  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS Ghaddar, G., Ruchon, A.F., Carpentier, M., Marcinkiewicz, M.,  
Seidah, N.G., Crine, P., Desrosiers, L. and Boileau, G.  
TITLE Molecular cloning and biochemical characterization of a new mouse  
testis soluble-zinc-metallopeptidase of the neprilysin family  
JOURNAL Biochem. J. 347 (Pt 2), 419-429 (2000)  
PUBMED 10749671  
REFERENCE 2 (bases 1 to 2925)  
AUTHORS Ghaddar, G., Ruchon, A.F., DesGroseillers, L. and Boileau, G.  
TITLE Direct Submission  
JOURNAL Submitted (09-AUG-1999) Biochemistry, University of Montreal, 2900  
Edouard Montpetit, Montreal, Que H3T 1J4, Canada

FEATURES  
source 1..2925  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/tissue\_type="testis"  
1..2925  
/gene="N11"  
332..2629  
/gene="N11"  
/note="N11; metallo-endopeptidase"  
/codon\_start=1  
/product="neprilysin-like metallopeptidase 1"  
/protein\_id="AAF69247.1"  
/db\_xref="GI:7769083"

/translation="MVRAGWCRKSPGVEYGLMVLVLLLLGAVTLGVFYSIGKQL  
PLTSLHFSWDETVVYKRALRDSLSKSDICTPSCVIAAARIENMDQSRNCFNY  
QYACGGWLRHVIPETNSRYSVFDILRDELEVILKGVLEDSTSOHRPAVERAKTYLRS  
CMQSVTEKRDSEPLSVLKMVGMPVAMDKWNETMGLKWELEQLAVLNSQFNRLV  
IDLFIWDDQNSRHVIYIDOPTLGMPSREYVFOEDNNHKVKAYLEFMTSVATMLRK  
DQNLKESAMVREMAEYVLETHLANATVPEKRHDVTALYHRMDLMELOERFGLKG  
FNWTLFTONVLSSEVEVLPDEEVVYVIGPYLENLEIDISYARTMQLVWRVLVD  
RIGLSQRFKEARVDYRKALYGTVEEVNRWCYVNSNMESAVGSLYIKRAFSDKS  
KSTVRELIEKIRSVFVDNLDELNMDESKKAKOAKMNIHQIGYDPYIILEDNNKHL  
DEEYSSLTFFEDLYFENGLONLKNNAQBSLKLREKVDONLWIGAAVVNAFYSPNRN  
QIVFPAGILOPPFFSKDQPSLNFSGGIGVIGHEITHGDFDNGRNFKNMGLDWSN  
FSARHFQOQSCMIYQGNFSEWELADNONVNGFSSLGENIADNGVROQAYKAYLRWLA  
DGGKDRPLGLNTYALQFFINAYQVWCGSYRPFVAVQSIKTDVHSPLKRYVLGSLQN  
LPGFSEAFHCPRGSPHMPMKRCRIW"

BASE COUNT 710 a 797 c 836 g 582 t  
ORIGIN

Query Match 91.6%; Score 17.4; DB 6; Length 2925;  
Best Local Similarity 94.7%; Pred. No. 1.5e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATCGGCTACCTGACTAC 19  
Db 1810 GATTGGCTACCTGACTAC 1828

RESULT 11  
AL607032  
LOCUS AL607032 208249 bp DNA linear HTG 24-JUL-2002  
DEFINITION Mus musculus chromosome 4 clone RP23-15L19, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, 2 unordered pieces.  
ACCESSION AL607032  
VERSION AL607032.15 GI:21955491  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_ACTIVEFIN; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE Mus musculus

LOCUS AF176569 2925 bp mRNA linear ROD 11-MAY-2000  
DEFINITION Mus musculus neprilysin-like metallopeptidase 1 (N11) mRNA,  
complete cds.  
ACCESSION AF176569  
VERSION AF176569.1 GI:7769082  
KEYWORDS  
SOURCE Mus musculus

ORGANISM  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS Ghaddar, G., Ruchon, A.F., Carpentier, M., Marcinkiewicz, M.,  
Seidah, N.G., Crine, P., Desrosiers, L. and Boileau, G.  
TITLE Molecular cloning and biochemical characterization of a new mouse  
testis soluble-zinc-metallopeptidase of the neprilysin family  
JOURNAL Biochem. J. 347 (Pt 2), 419-429 (2000)  
PUBMED 10749671  
REFERENCE 2 (bases 1 to 2925)  
AUTHORS Ghaddar, G., Ruchon, A.F., DesGroseillers, L. and Boileau, G.  
TITLE Direct Submission  
JOURNAL Submitted (09-AUG-1999) Biochemistry, University of Montreal, 2900  
Edouard Montpetit, Montreal, Que H3T 1J4, Canada

FEATURES  
source 1..2925  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/tissue\_type="testis"  
1..2925  
/gene="N11"  
332..2629  
/gene="N11"  
/note="N11; metallo-endopeptidase"  
/codon\_start=1  
/product="neprilysin-like metallopeptidase 1"  
/protein\_id="AAF69247.1"  
/db\_xref="GI:7769083"

/translation="MVRAGWCRKSPGVEYGLMVLVLLLLGAVTLGVFYSIGKQL  
PLTSLHFSWDETVVYKRALRDSLSKSDICTPSCVIAAARIENMDQSRNCFNY  
QYACGGWLRHVIPETNSRYSVFDILRDELEVILKGVLEDSTSOHRPAVERAKTYLRS  
CMQSVTEKRDSEPLSVLKMVGMPVAMDKWNETMGLKWELEQLAVLNSQFNRLV  
IDLFIWDDQNSRHVIYIDOPTLGMPSREYVFOEDNNHKVKAYLEFMTSVATMLRK  
DQNLKESAMVREMAEYVLETHLANATVPEKRHDVTALYHRMDLMELOERFGLKG  
FNWTLFTONVLSSEVEVLPDEEVVYVIGPYLENLEIDISYARTMQLVWRVLVD  
RIGLSQRFKEARVDYRKALYGTVEEVNRWCYVNSNMESAVGSLYIKRAFSDKS  
KSTVRELIEKIRSVFVDNLDELNMDESKKAKOAKMNIHQIGYDPYIILEDNNKHL  
DEEYSSLTFFEDLYFENGLONLKNNAQBSLKLREKVDONLWIGAAVVNAFYSPNRN  
QIVFPAGILOPPFFSKDQPSLNFSGGIGVIGHEITHGDFDNGRNFKNMGLDWSN  
FSARHFQOQSCMIYQGNFSEWELADNONVNGFSSLGENIADNGVROQAYKAYLRWLA  
DGGKDRPLGLNTYALQFFINAYQVWCGSYRPFVAVQSIKTDVHSPLKRYVLGSLQN  
LPGFSEAFHCPRGSPHMPMKRCRIW"

BASE COUNT 710 a 797 c 836 g 582 t  
ORIGIN

Query Match 91.6%; Score 17.4; DB 10; Length 2925;  
Best Local Similarity 94.7%; Pred. No. 1.5e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATCGGCTACCTGACTAC 19  
Db 1810 GATTGGCTACCTGACTAC 1828

RESULT 11  
AL607032  
LOCUS AL607032 208249 bp DNA linear HTG 24-JUL-2002  
DEFINITION Mus musculus chromosome 4 clone RP23-15L19, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, 2 unordered pieces.  
ACCESSION AL607032  
VERSION AL607032.15 GI:21955491  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_ACTIVEFIN; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE Mus musculus

ORGANISM  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS Ghaddar, G., Ruchon, A.F., Carpentier, M., Marcinkiewicz, M.,  
Seidah, N.G., Crine, P., Desrosiers, L. and Boileau, G.  
TITLE Molecular cloning and biochemical characterization of a new mouse  
testis soluble-zinc-metallopeptidase of the neprilysin family  
JOURNAL Biochem. J. 347 (Pt 2), 419-429 (2000)  
PUBMED 10749671  
REFERENCE 2 (bases 1 to 2925)  
AUTHORS Ghaddar, G., Ruchon, A.F., DesGroseillers, L. and Boileau, G.  
TITLE Direct Submission  
JOURNAL Submitted (09-AUG-1999) Biochemistry, University of Montreal, 2900  
Edouard Montpetit, Montreal, Que H3T 1J4, Canada

FEATURES  
source 1..2925  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/tissue\_type="testis"  
1..2925  
/gene="N11"  
332..2629  
/gene="N11"  
/note="N11; metallo-endopeptidase"  
/codon\_start=1  
/product="neprilysin-like metallopeptidase 1"  
/protein\_id="AAF69247.1"  
/db\_xref="GI:7769083"

/translation="MVRAGWCRKSPGVEYGLMVLVLLLLGAVTLGVFYSIGKQL  
PLTSLHFSWDETVVYKRALRDSLSKSDICTPSCVIAAARIENMDQSRNCFNY  
QYACGGWLRHVIPETNSRYSVFDILRDELEVILKGVLEDSTSOHRPAVERAKTYLRS  
CMQSVTEKRDSEPLSVLKMVGMPVAMDKWNETMGLKWELEQLAVLNSQFNRLV  
IDLFIWDDQNSRHVIYIDOPTLGMPSREYVFOEDNNHKVKAYLEFMTSVATMLRK  
DQNLKESAMVREMAEYVLETHLANATVPEKRHDVTALYHRMDLMELOERFGLKG  
FNWTLFTONVLSSEVEVLPDEEVVYVIGPYLENLEIDISYARTMQLVWRVLVD  
RIGLSQRFKEARVDYRKALYGTVEEVNRWCYVNSNMESAVGSLYIKRAFSDKS  
KSTVRELIEKIRSVFVDNLDELNMDESKKAKOAKMNIHQIGYDPYIILEDNNKHL  
DEEYSSLTFFEDLYFENGLONLKNNAQBSLKLREKVDONLWIGAAVVNAFYSPNRN  
QIVFPAGILOPPFFSKDQPSLNFSGGIGVIGHEITHGDFDNGRNFKNMGLDWSN  
FSARHFQOQSCMIYQGNFSEWELADNONVNGFSSLGENIADNGVROQAYKAYLRWLA  
DGGKDRPLGLNTYALQFFINAYQVWCGSYRPFVAVQSIKTDVHSPLKRYVLGSLQN  
LPGFSEAFHCPRGSPHMPMKRCRIW"

BASE COUNT 710 a 797 c 836 g 582 t  
ORIGIN

Query Match 91.6%; Score 17.4; DB 10; Length 2925;  
Best Local Similarity 94.7%; Pred. No. 1.5e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATCGGCTACCTGACTAC 19  
Db 1810 GATTGGCTACCTGACTAC 1828

RESULT 11  
AL607032  
LOCUS AL607032 208249 bp DNA linear HTG 24-JUL-2002  
DEFINITION Mus musculus chromosome 4 clone RP23-15L19, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, 2 unordered pieces.  
ACCESSION AL607032  
VERSION AL607032.15 GI:21955491  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_ACTIVEFIN; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE Mus musculus

ORGANISM  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS Ghaddar, G., Ruchon, A.F., Carpentier, M., Marcinkiewicz, M.,  
Seidah, N.G., Crine, P., Desrosiers, L. and Boileau, G.  
TITLE Molecular cloning and biochemical characterization of a new mouse  
testis soluble-zinc-metallopeptidase of the neprilysin family  
JOURNAL Biochem. J. 347 (Pt 2), 419-429 (2000)  
PUBMED 10749671  
REFERENCE 2 (bases 1 to 2925)  
AUTHORS Ghaddar, G., Ruchon, A.F., DesGroseillers, L. and Boileau, G.  
TITLE Direct Submission  
JOURNAL Submitted (09-AUG-1999) Biochemistry, University of Montreal, 2900  
Edouard Montpetit, Montreal, Que H3T 1J4, Canada

FEATURES  
source 1..2925  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/tissue\_type="testis"  
1..2925  
/gene="N11"  
332..2629  
/gene="N11"  
/note="N11; metallo-endopeptidase"  
/codon\_start=1  
/product="neprilysin-like metallopeptidase 1"  
/protein\_id="AAF69247.1"  
/db\_xref="GI:7769083"

/translation="MVRAGWCRKSPGVEYGLMVLVLLLLGAVTLGVFYSIGKQL  
PLTSLHFSWDETVVYKRALRDSLSKSDICTPSCVIAAARIENMDQSRNCFNY  
QYACGGWLRHVIPETNSRYSVFDILRDELEVILKGVLEDSTSOHRPAVERAKTYLRS  
CMQSVTEKRDSEPLSVLKMVGMPVAMDKWNETMGLKWELEQLAVLNSQFNRLV  
IDLFIWDDQNSRHVIYIDOPTLGMPSREYVFOEDNNHKVKAYLEFMTSVATMLRK  
DQNLKESAMVREMAEYVLETHLANATVPEKRHDVTALYHRMDLMELOERFGLKG  
FNWTLFTONVLSSEVEVLPDEEVVYVIGPYLENLEIDISYARTMQLVWRVLVD  
RIGLSQRFKEARVDYRKALYGTVEEVNRWCYVNSNMESAVGSLYIKRAFSDKS  
KSTVRELIEKIRSVFVDNLDELNMDESKKAKOAKMNIHQIGYDPYIILEDNNKHL  
DEEYSSLTFFEDLYFENGLONLKNNAQBSLKLREKVDONLWIGAAVVNAFYSPNRN  
QIVFPAGILOPPFFSKDQPSLNFSGGIGVIGHEITHGDFDNGRNFKNMGLDWSN  
FSARHFQOQSCMIYQGNFSEWELADNONVNGFSSLGENIADNGVROQAYKAYLRWLA  
DGGKDRPLGLNTYALQFFINAYQVWCGSYRPFVAVQSIKTDVHSPLKRYVLGSLQN  
LPGFSEAFHCPRGSPHMPMKRCRIW"





```

misc_feature 8192..34992
/note="assembly_name:Contig183"
misc_feature 35093..109409
/note="assembly_name:Contig184"
misc_feature 109510..233032
/note="assembly_name:Contig185"
misc_feature 233133..233265
/note="assembly_name:Contig18"
misc_feature 233366..233488
/note="assembly_name:Contig19"
BASE COUNT 62961 a 57044 c 53747 g 59136 t 600 others
ORIGIN

```

```

Query Match 89.5%; Score 17; DB 2; Length 233488;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 TCGGCTACCTGACTAC 19
|||||
Db 27604 TCGGCTACCTGACTAC 27588
|||||

```

```

RESULT 13
AC090432/c 193363 bp DNA linear HTG 20-FEB-2002
LOCUS Mus musculus chromosome 5 clone RP23-454P8 strain C57BL6/J, WORKING
DEFINITION DRAFT SEQUENCE, 7 ordered pieces.
ACCESSION AC090432
VERSION AC090432.2 GI:18767482
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Mus musculus.

```

```

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 193363)
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Hachichi, P.,
Ho, S.-L., Idoi, J.R., Karlins, E., Laric, P., Lee-Lin, S.-Q.,
Legaspi, R., Maduro, Q.L., Maduro, V.B., Masiello, C., Maskeri, B.,
Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A.,
Stantripp, S., Thomas, J.W., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

```

```

TITLE NISC Comparative Sequencing Initiative
REFERENCE 2 (bases 1 to 193363)
AUTHORS Green, E.D.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2001) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3 (bases 1 to 193363)
AUTHORS Green, E.D.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
COMMENT On Feb 20, 2002 this sequence version replaced gi:13027532.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.mouse@hgrl.nih.gov
----- Project Information
Center project name: gn
Center clone name: 454P08

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition,

the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 192013 bases at least Q40  
 Consensus quality: 192474 bases at least Q30  
 Consensus quality: 192660 bases at least Q20  
 Insert size: 179000; agarose-fp  
 Insert size: 209000; pulse-field-gel  
 Insert size: 192763; sum-of-contigs  
 Quality coverage: 13.04x in Q20 bases; agarose-fp  
 Quality coverage: 11.17x in Q20 bases; pulse-field-gel  
 Quality coverage: 12.11x in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 7 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* been provided by the submitter.

\* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

```

1 18890: contig of 18890 bp in length
* 18891 18990: gap of unknown length
* 18991 48374: contig of 29384 bp in length
* 48375 48474: gap of unknown length
* 48475 112619: contig of 64145 bp in length
* 112620 112719: gap of unknown length
* 112720 116063: contig of 3344 bp in length
* 116064 116163: gap of unknown length
* 116164 123841: contig of 7678 bp in length
* 123842 123941: gap of unknown length
* 123942 187728: contig of 63787 bp in length
* 187729 187828: gap of unknown length
* 187829 193363: contig of 5535 bp in length.

```

#### FEATURES

```

source
1..193363
/organism="Mus musculus"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="5"
/clone="RP23-454P8"
/clone_lib="RPCI mouse BAC library 23"
1..30346
/note="clone overlaps with GenBank Accession Number
AC090443 clone RP23-434H17 (center project name gm)"
misc_feature
1..30346
/note="assembly_fragment
clone end:SP6
vector_side:left"
misc_feature
18991..48374
/note="assembly_fragment"
48475..112619
/note="assembly_fragment"
112720..116063
/note="assembly_fragment"
116164..123841
/note="assembly_fragment"
123942..187728
/note="assembly_fragment"
187729..193363
/note="assembly_fragment"
AC026478 clone RP23-135F23 (center project name ye)"
187829..193363
/note="assembly_fragment
clone end:T7
vector_side:right"

```

```

BASE COUNT 57206 a 42427 c 41080 g 52049 t 601 others
ORIGIN
Query Match 86.3%; Score 16.4; DB 2; Length 193363;
Best Local Similarity 94.4%; Pred. No. 3.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATCGGCTACCTGACTA 18
|||||
Db 160167 GATCGGCTACCTGACTA 160150

RESULT 14
AC121567
LOCUS
DEFINITION Mus musculus chromosome UNK clone RP23-225D14, WORKING DRAFT.
ACCESSION AC121567
VERSION AC121567.1 GI:20986638
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1. (bases 1 to 204867)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2. (bases 1 to 204867)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (20-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

REFERENCE 3. (bases 1 to 204867)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

COMMENT
----- Genome Center
Center: Washington University Genome Sequencing Center
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M_LBA0225D14
----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-terminator Big Dye; 100%
Assembly: Dye-terminator Big Dye; 100% of reads
Consensus quality: 203471 bases at least Q40
Consensus quality: 203858 bases at least Q30
Consensus quality: 203980 bases at least Q20
Insert size: 136000; agarose-1p
Insert size: 205765; sum-of-contigs
Quality coverage: 12.50 in Q20 bases; agarose-1p
Quality coverage: 12.91 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 5227: contig of 5227 bp in length
* 5328: gap of unknown length
* 28642: contig of 23315 bp in length
* 28643: gap of unknown length

```

---

```

* 28743 104070: contig of 75328 bp in length
* 104071 104170: gap of unknown length
* 104171 204867: contig of 100697 bp in length.
FEATURES
source
1. 204867
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-225D14"
misc_feature
1. 5227
/notes="assembly_name:Contig5"
misc_feature
5328..28642
/notes="assembly_name:Contig6"
misc_feature
28743..104070
/notes="assembly_name:Contig7"
misc_feature
104171..204867
/notes="assembly_name:Contig8"
BASE COUNT 65519 a 37283 c 36747 g 65012 t 306 others
ORIGIN
Query Match 86.3%; Score 16.4; DB 2; Length 204867;
Best Local Similarity 94.4%; Pred. No. 3.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATCGGCTACCTGACTAC 19
|||||
Db 127968 ATAGGCTACCTGACTAC 127985

RESULT 15
AC115760
LOCUS
DEFINITION Mus musculus clone RP23-62M14, WORKING DRAFT SEQUENCE, 18 ordered
pieces.
ACCESSION AC115760
VERSION AC115760.2 GI:21592113
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1. (bases 1 to 206132)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Unpublished
JOURNAL Submitted
REFERENCE 2. (bases 1 to 206132)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Govette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R.,
Landers,T., Lehotzky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meidrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Olliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
JOURNAL Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3. (bases 1 to 206132)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

```

Anderson, S., Barna, N., Bastien, V., Bloom, T., Boquslavskiy, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R., Landers, T., Lehoczyk, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneses, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 26, 2002 this sequence version replaced gi:19683412.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RN/RepeatMasker.html>

TITLE  
JOURNAL  
COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: 62\_M\_14

Center clone name: 62\_M\_14

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 199666 bases at least Q40

Consensus quality: 202898 bases at least Q30

Consensus quality: 203843 bases at least Q20

Insert size: 210000; agarose-fp

Insert size: 204432; sum-of-contigs

Quality coverage: 7.1 in Q20 bases; agarose-fp

Quality coverage: 7.3 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 18 contigs. Gaps between the contigs

\* are represented as runs of N. The order of the pieces

\* is believed to be correct as given, however the sizes

\* of the gaps between them are based on estimates that have

\* provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1 3890: contig of 3890 bp in length.

\* 3891 3990: gap of 100 bp

\* 3991 80283: contig of 76293 bp in length

\* 80284 80383: gap of 100 bp

\* 80384 81288: contig of 905 bp in length

\* 81289 81388: gap of 100 bp

\* 81389 83210: contig of 1822 bp in length

\* 83211 83310: gap of 100 bp

\* 83311 84999: contig of 1689 bp in length

\* 85000 85099: gap of 100 bp

\* 85100 86992: contig of 1893 bp in length

\* 86993 87092: gap of 100 bp

\* 87093 88392: contig of 1300 bp in length

\* 88393 88492: gap of 100 bp

\* 88493 90835: contig of 2343 bp in length

\* 90836 90935: gap of 100 bp

\* 90936 93492: contig of 2557 bp in length

\* 93493 93592: gap of 100 bp

\* 93593 96059: contig of 2467 bp in length

\* 96060 96159: gap of 100 bp

\* 96160 97912: contig of 1753 bp in length

\* 97913 98012: gap of 100 bp

\* 98013 99897: contig of 1885 bp in length

\* 99898 99997: gap of 100 bp

\* 99998 103113: contig of 3116 bp in length

\* 103114 103213: gap of 100 bp

\* 103214 107568: contig of 4355 bp in length

\* 107569 107668: gap of 100 bp

\* 107669 116549: contig of 8881 bp in length

\* 116550 116649: gap of 100 bp

\* 116650 128482: contig of 11833 bp in length

\* 128483 128582: gap of 100 bp

\* 128583 205750: contig of 77168 bp in length

\* 205751 205850: gap of 100 bp

\* 205851 206132: contig of 282 bp in length.

# FEATURES

## Source

1. 206132

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="RP23-62M14"

/clone\_lib="RPCI-23 Female Mouse BAC"

1. 3890

/note="assembly\_fragment"

clone\_end:SP6

vector\_side:left

3991. 80283

/note="assembly\_fragment"

80384. 81288

/note="assembly\_fragment"

81389. 83210

/note="assembly\_fragment"

83311. 84999

/note="assembly\_fragment"

85100. 86992

/note="assembly\_fragment"

87093. 88392

/note="assembly\_fragment"

88493. 90835

/note="assembly\_fragment"

90936. 93492

/note="assembly\_fragment"

93593. 96059

/note="assembly\_fragment"

96160. 97912

/note="assembly\_fragment"

98013. 99897

/note="assembly\_fragment"

99998. 103113

/note="assembly\_fragment"

103214. 107568

/note="assembly\_fragment"

107669. 116549

/note="assembly\_fragment"

116650. 128482

/note="assembly\_fragment"

128583. 205750

/note="assembly\_fragment"

205851. 206132

/note="assembly\_fragment"

clone\_end:T7

vector\_side:right

BASE COUNT 58333 a 44062 c 44541 g 57492 t 1704 others

## ORIGIN

Query Match 86.3%; Score 16.4; DB 2; Length 206132;

Best Local Similarity 94.4%; Pred. No. 3.4e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATCGGCTACCTGACTAC 19

Db 202676 ATCGGTTACCGTGACTAC 202693  
||||| |||||||||

Search completed: July 8, 2003, 03:35:08  
Job time : 211.093 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 00:43:28 ; Search time 119.376 Seconds  
(without alignments)  
358.431 Million cell updates/sec

Title: US-09-647-780A-19

Perfect score: 19

Sequence: 1 gtcggctacctgactac 19

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	100.0	19	AAZ28825	Rat membrane metal
2	19	100.0	2286	AAZ28130	Soluble secreted e
3	19	100.0	2765	AAZ28810	Rat membrane metal
c 4	17.4	91.6	29	AAAG3760	PCR primer used to
5	17.4	91.6	2925	AAAG3763	CDNA encoding neut
c 6	15.8	83.2	1422	AAAG3763	DNA encoding novel
c 7	15.8	83.2	1497	AAAG3763	DNA encoding novel
8	15.8	83.2	2076	AAAG3737	Nucleotide sequenc
9	15.8	83.2	2232	AAAG37186	Human metalloprote

10	15.8	83.2	2262	AAAG3739	Nucleotide sequenc
11	15.8	83.2	2318	AAAG30580	Human protease, PR
12	15.8	83.2	2340	AAAG3738	Nucleotide sequenc
c 13	15.8	83.2	2529	ABL222299	Drosophila melanog
14	15.8	83.2	2580	ABL222299	Human SEP endopept
15	15.8	83.2	2636	AAAG3760	Human neprilysin-1
16	15.8	83.2	2663	AAAG3761	Human neprilysin-1
17	15.8	83.2	2676	AAAG3764	CDNA encoding neut
18	15.8	83.2	2714	AAAG3765	Human neprilysin-1
19	15.8	83.2	2893	AAAG3769	Human SEP endopept
20	15.8	83.2	2893	AAAG3769	Human soluble secr
c 21	15.8	83.2	2928	AAAG3769	DNA encoding novel
22	15.8	83.2	2953	AAAG3769	Human SEP CDNA inc
23	15.8	83.2	2975	AAAG3769	Drosophila melanog
c 24	15.8	83.2	30610	ABL222298	Drosophila melanog
c 25	15.8	83.2	30610	ABL222298	Human nervous syst
26	15.8	83.2	30610	ABL222298	Human polynucleoti
c 27	15.4	81.1	351	AAAG3769	Arabidopsis thalia
28	15.4	81.1	483	AAAG3769	Sequence #65 used
29	15.4	81.1	574	AAAG3769	Plant dwarfing/stu
30	15.4	81.1	574	AAAG3769	Arabidopsis thalia
c 31	15.4	81.1	585	AAAG3769	Plant dwarfing/stu
c 32	15.4	81.1	613	AAAG3769	A. thaliana psaeI
c 33	15.4	81.1	676	AAAG3769	Arabidopsis thalia
c 34	15.4	81.1	967	AAAG3769	Sequence #65 used
35	15	78.9	225	AAAG3769	Non-B, non-C, non-
c 36	14.8	77.9	306	AAAG3769	Corn tassell-deri
c 37	14.8	77.9	634	AAAG3769	DNA encoding novel
c 38	14.8	77.9	2635	AAAG3769	Drosophila melanog
39	14.8	77.9	5250	AAAG3769	Genomic sequence #
c 40	14.8	77.9	5250	AAAG3769	Human reproductive
41	14.8	77.9	5276	AAAG3769	Drosophila melanog
42	14.8	77.9	5987	AAAG3769	Nucleotide sequenc
43	14.8	77.9	7397	AAAG3769	Drosophila melanog
c 44	14.4	75.8	337	AAAG3769	Human OREFX polynuc
45	14.4	75.8	1737	AAAG3769	Drosophila melanog

ALIGNMENTS

RESULT 1

AAZ228825

ID AAZ228825 standard; DNA; 19 BP.

XX AAZ228825;

XX AAZ228825;

DT 01-FEB-2000 (first entry)

XX 01-FEB-2000 (first entry)

DE Rat membrane metalloprotease NEPII gene probe #15.

XX Rat; membrane metalloprotease; neprilysin II; NEPII; inactivation; ss;  
KW neuron; hormone; peptide messenger; inhibitor; detection; disorder;  
KW cardiovascular disease; neurodegenerative disease; growth disorder;  
KW hypothalamic-hypophyseal axis; endocrine disorder; probe; hybridisation.

OS Synthetic.

OS Rattus rattus.

XX FR2777291-A1.

XX 15-OCT-1999.

XX 08-APR-1998; 98FR-0004389.

XX 08-APR-1998; 98FR-0004389.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Oulmet T, Gros C, Haret C, Bonhomme MC, Facchinetti P;

XX Schwartz JG;

XX WPI; 1999-593429/51.

XX New membrane metalloprotease NEP II, involved in proteolysis of  
PT neuronal and hormonal peptides, used to screen for inhibitors,  
PT potentially useful for treating e.g. cardiovascular disease  
XX  
PS Claim 3; Page 24; 29pp; French.  
XX  
CC Sequences AA228811-228827 represent probes for detecting the rat  
CC membrane metalloprotease designated neprilysine II (NEPII) gene  
CC (AA228810). NEPII is involved in (in)activation of neuronal and hormonal  
CC peptide messengers. NEPII is used to screen for specific substrates (used  
CC to detect NEPII in cells and tissues) or inhibitors, which can also be  
CC used to detect NEPII or for treatment of disorders related to peptidergic  
CC signalling in which NEPII is involved, e.g. cardiovascular or  
CC neurodegenerative diseases; growth disorders of endocrine origin;  
CC disturbances of the hypothalamic- hypophyseal axis or endocrine  
CC disorders.  
XX  
SQ Sequence 19 BP; 4 A; 7 C; 4 G; 4 T; 0 other;  
Query Match 100.0%; Score 19; DB 20; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 GATCGGCTACCTGACTAC 19  
DB 1 GATCGGCTACCTGACTAC 19  
RESULT 2  
AAD28130  
ID AAD28130 standard; DNA; 2286 BP.  
XX  
AC AAD28130;  
XX  
DT 07-MAY-2002 (first entry).  
XX  
DE Soluble secreted endopeptidase (SEP) consensus DNA.  
XX  
KW Human; secreted endopeptidase; SEP; endocrine; vasotropic; analgesic;  
KW enzyme; gynaecological; antisense-therapy; male erectile dysfunction;  
KW MED; female sexual dysfunction; FSD; female sexual arousal disorder;  
KW FSAD; premature ejaculation; anorgasmia; vaginismus; mouse; rat; ds.  
XX  
OS Homo sapiens.  
OS Mus sp.  
OS Rattus sp.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 1664..2286  
FT /\*tag= a  
FT /\*note= "Encodes catalytic domain"  
XX  
PN WO200206492-A1.  
XX  
PD 24-JAN-2002.  
XX  
PF 16-JUL-2001; 2001WO-IB01263.  
XX  
PR 14-JUL-2000; 2000GB-0017387.  
XX  
XX (PFIZ ) PFIZER LTD.  
PA (PFIZ ) PFIZER INC.  
XX  
XX Harrow ID, Stacey P, Walsh RT, Wayman CP, Wayman CP, Phillips SC;  
XX WPI; 2002-155042/20.  
XX  
PT An isolated and/or purified nucleic acid encoding a human soluble  
PT secreted endopeptidase which is useful for treating sexual dysfunction,  
PT for e.g. male erectile dysfunction or female sexual dysfunction such as  
PT female sexual arousal disorder  
XX

PS Disclosure; Fig 6; 167pp; English.  
XX  
CC The invention relates to an isolated and/or purified nucleic acid  
CC encoding a human soluble secreted endopeptidase (SEP). The anti-SEP  
CC antibody and the compound which inhibits or selectively inhibits the  
CC human SEP protein are useful in the manufacture of a medicament for  
CC the prophylaxis and/or treatment of sexual dysfunction, in particular  
CC male erectile dysfunction (MED) or female sexual dysfunction (FSD)  
CC (e.g. female sexual arousal disorder (FSAD)). The SEP polypeptides are  
CC also useful for treating the above disorders and other disorders such  
CC as premature ejaculation, anorgasmia, sexual pain (e.g. vaginismus) and  
CC hypoaffective sexual desire disorder. The present sequence is SEP consensus  
CC DNA sequence found in human, mouse and rat.  
XX  
SQ Sequence 2286 BP; 565 A; 601 C; 674 G; 446 T; 0 other;  
Query Match 100.0%; Score 19; DB 24; Length 2286;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 GATCGGCTACCTGACTAC 19  
DB 1477 GATCGGCTACCTGACTAC 1495  
RESULT 3  
AAZ28810  
ID AAZ28810 standard; cDNA; 2765 BP.  
XX  
AC AAZ28810;  
XX  
DT 01-FEB-2000 (first entry)  
XX  
DE Rat membrane metalloprotease NEPII gene.  
XX  
KW Rat; membrane metalloprotease; neprilysine II; NEPII; inactivation;  
KW neuron; hormone; peptide messenger; inhibitor; detection; disorder;  
KW cardiovascular disease; neurodegenerative disease; growth disorder;  
KW hypothalamic-hypophyseal axis; endocrine disorder; ds.  
XX  
OS Rattus rattus.  
XX  
PN FR2777291-A1.  
XX  
PD 15-OCT-1999.  
XX  
PF 08-APR-1998; 98FR-0004389.  
XX  
PR 08-APR-1998; 98FR-0004389.  
XX  
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
PA  
PI Oulmet T, Gros C, Haret C, Bonhomme MC, Facchinetti P;  
PI Schwartz JC;  
XX  
XX WPI; 1999-593429/51.  
DR P-PSDB; AAY4177.  
XX  
XX New membrane metalloprotease NEP II, involved in proteolysis of  
PT neuronal and hormonal peptides, used to screen for inhibitors,  
PT potentially useful for treating e.g. cardiovascular disease  
XX  
PS Claim 2; Page 12-16; 29pp; French.  
XX  
CC This sequence represents the gene for the rat membrane metalloprotease  
CC designated neprilysine II (NEPII), which is involved in (in)activation  
CC of neuronal and hormonal peptide messengers. NEPII is used to screen  
CC for specific substrates (used to detect NEPII in cells and tissues) or  
CC inhibitors, which can also be used to detect NEPII or for treatment of  
CC disorders related to peptidergic signalling in which NEPII is involved,  
CC e.g. cardiovascular or neurodegenerative diseases; growth disorders of  
CC endocrine origin; disturbances of the hypothalamic-hypophyseal axis or  
CC endocrine disorders.

```

XX SQ Sequence 2765 BP; 684 A; 735 C; 787 G; 559 T; 0 other;
Query Match 100.08; Score 19; DB 20; Length 2765;
Best Local Similarity 100.08; Pred. No. 2.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCGGCTACCTGACTAC 19
DB 1612 GATCGGCTACCTGACTAC 1630

RESULT 4
AAA63760/c
ID AAA63760 standard; DNA; 29 BP.
XX AC AAA63760;
XX 04-DEC-2000 (first entry)
XX DT
XX DE
XX NEprilysin; neutral endopeptidase metalloproteinase-like enzyme;
KW NEP-like enzyme; protein production; protein secretion;
KW neurological disease; Alzheimer's disease; pain; psychiatric disorder;
KW fertility; bone disease; abnormal phosphate metabolism; PCR primer; ss.
XX OS
XX Mus sp.
XX WO2000047750-A2.
XX 17-AUG-2000.
XX 11-FEB-2000; 2000WO-CA00147.
XX 11-FEB-1999; 99CA-2260376.
XX (UYMO-) UNIV MONTREAL.
XX Desgroseillers L, Boileau G;
XX WPI; 2000-549148/50.
XX Novel neutral endopeptidase-like metalloproteinase polypeptides and
PT polynucleotides, used to screen for related sequences and enzyme
PT inhibitors, used for the treatment of NL-3 related bone disorders -
XX Disclosure; Page 14; 59pp; English.
XX PCR primers AAA63759-60 were used to amplify a murine cDNA sequence
CC encoding a neutral endopeptidase metalloproteinase (NEP)-like
CC polypeptide, designated NL-1. The specification also describes NL-2 and
CC NL-3. The NL enzymes are used to test for specific inhibitors. The
CC N-terminal region of the enzymes can be used to promote production and
CC secretion of foreign proteins and active biopeptides, using chimeric
CC constructs containing the foreign protein downstream from and in phase
CC with the N-terminal region. The NL enzymes have been localised to
CC the brain, and may be useful in the treatment of neurological diseases
CC such as Alzheimer's disease, pain, and psychiatric disorders. NL enzymes
CC have also been localised to the testis and ovaries, and may be used to
CC control fertility. They have also been localised to bones, and may be
CC used to treat bone diseases, and abnormal phosphate metabolisms related
CC to improper peptide processing by the NL-3 enzyme.
XX SQ Sequence 29 BP; 9 A; 7 C; 6 G; 7 T; 0 other;
Query Match 91.68; Score 17.4; DB 21; Length 29;
Best Local Similarity 94.7%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATCGGCTACCTGACTAC 19
DB 29 GATCGGCTACCTGACTAC 11

```

## RESULT 5

AAA63763  
ID AAA63763 standard; cDNA; 2925 BP.

XX AC AAA63763;

XX 04-DEC-2000 (first entry)

XX cDNA encoding neutral endopeptidase metalloproteinase-like enzyme NL-1.

XX NEprilysin; neutral endopeptidase metalloproteinase-like enzyme;  
KW NEP-like enzyme; protein production; protein secretion;  
KW neurological disease; Alzheimer's disease; pain; psychiatric disorder;  
KW fertility; bone disease; abnormal phosphate metabolism; ss.

XX OS

XX Mus sp.

XX Key Location/Qualifiers  
XX CDS 332..2629  
XX FT /\*tag= a  
XX FT /product= "neutral endopeptidase metalloproteinase-like  
XX enzyme NL-1"

XX PN WO2000047750-A2.

XX 17-AUG-2000.

XX 11-FEB-2000; 2000WO-CA00147.

XX 11-FEB-1999; 99CA-2260376.

XX (UYMO-) UNIV MONTREAL.

XX Desgroseillers L, Boileau G;

XX WPI; 2000-549148/50.

XX P-PSDB; AAB08130.

XX Novel neutral endopeptidase-like metalloproteinase polypeptides and  
PT polynucleotides, used to screen for related sequences and enzyme  
PT inhibitors, used for the treatment of NL-3 related bone disorders.

XX Disclosure; Fig 3; 59pp; English.

XX The present sequence encodes a murine neutral endopeptidase  
CC metalloproteinase-like enzyme, designated NL-1. The specification  
CC also describes NL-2 and NL-3. The NL enzymes are used to test for  
CC specific inhibitors. The N-terminal region of the enzymes can be used  
CC to promote production and secretion of foreign proteins and active  
CC biopeptides, using chimeric constructs containing the foreign protein  
CC downstream from and in phase with the N-terminal region. The NL enzymes  
CC are have been localised to the brain, and may be useful in the  
CC treatment of neurological diseases such as Alzheimer's disease, pain,  
CC and psychiatric disorders. NL enzymes have also been localised to the  
CC testis and ovaries, and may be used to control fertility. They have  
CC also been localised to bones, and may be used to treat bone diseases,  
CC and abnormal phosphate metabolisms related to improper peptide  
CC processing by the NL-3 enzyme.

XX SQ Sequence 2925 BP; 710 A; 797 C; 836 G; 582 T; 0 other;

Query Match 91.6%; Score 17.4; DB 21; Length 2925;  
Best Local Similarity 94.7%; Pred. No. 19;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATCGGCTACCTGACTAC 19

DB 1810 GATCGGCTACCTGACTAC 1828

## RESULT 6.



AAS73814/c  
ID AAS73814 standard; cDNA: 1422 BP.

XX AC AAS73814;  
XX DT 13-FEB-2002. (first entry)  
XX DE DNA encoding novel human diagnostic protein #9618.  
XX DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX OS Homo sapiens.  
XX PN WO200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US08631.  
XX PR 31-MAR-2000; 2000US-0540217.  
XX PR 23-AUG-2000; 2000US-0649167.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Drmanac RT, Liu C, Tang YT;  
XX DR WPI: 2001-639362/73.  
XX DR P-PSDB; ABG09627.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
XX PT diagnostics, forensics, gene mapping, identification of mutations  
XX PT responsible for genetic disorders or other traits and to assess  
XX PT biodiversity.

XX PS Claim 1: SEQ ID No 9618; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and  
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX CC and gene mapping, and in recombinant production of (II). The  
XX CC polynucleotides are also used in diagnostics as expressed sequence tags  
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques  
XX CC to restore normal activity of (II) or to treat disease states involving  
XX CC (II). (II) is useful for generating antibodies against it, detecting or  
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as  
XX CC a food supplement. (II) and its binding partners are useful in medical  
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating  
XX CC disorders involving aberrant protein expression or biological activity.  
XX CC The polypeptide and polynucleotide sequences have applications in  
XX CC diagnostics, forensics, gene mapping, identification of mutations  
XX CC responsible for genetic disorders or other traits to assess biodiversity  
XX CC and to produce other types of data and products dependent on DNA and  
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human  
XX CC diagnostic coding sequences of the invention.  
XX CC Note: The sequence data for this patent did not appear in the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 1422 BP; 385 A; 367 C; 345 G; 325 T; 0 other;

Query Match 83.2%; Score 15.8; DB 23; Length 1422;  
Best Local Similarity 89.5%; Pred No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GATCGGCTACCTGACTAC 19  
||| ||||| ||||| |||||  
Db 22 GACCGGCTTCCTGACTAC 4

RESULT 7  
AAS69914/c  
ID AAS69914 standard; cDNA: 1497 BP.

XX AC AAS69914;  
XX DT 13-FEB-2002 (first entry)  
XX DE DNA encoding novel human diagnostic protein #5718.  
XX DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX OS Homo sapiens.  
XX PN WO200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US08631.  
XX PR 31-MAR-2000; 2000US-0540217.  
XX PR 23-AUG-2000; 2000US-0649167.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Drmanac RT, Liu C, Tang YT;  
XX DR WPI: 2001-639362/73.  
XX DR P-PSDB; ABG05727.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
XX PT diagnostics, forensics, gene mapping, identification of mutations  
XX PT responsible for genetic disorders or other traits and to assess  
XX PT biodiversity.

XX PS Claim 1: SEQ ID No 5718; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and  
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX CC and gene mapping, and in recombinant production of (II). The  
XX CC polynucleotides are also used in diagnostics as expressed sequence tags  
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques  
XX CC to restore normal activity of (II) or to treat disease states involving  
XX CC (II). (II) is useful for generating antibodies against it, detecting or  
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as  
XX CC a food supplement. (II) and its binding partners are useful in medical  
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating  
XX CC disorders involving aberrant protein expression or biological activity.  
XX CC The polypeptide and polynucleotide sequences have applications in  
XX CC diagnostics, forensics, gene mapping, identification of mutations  
XX CC responsible for genetic disorders or other traits to assess biodiversity  
XX CC and to produce other types of data and products dependent on DNA and  
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human  
XX CC diagnostic coding sequences of the invention.  
XX CC Note: The sequence data for this patent did not appear in the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 1497 BP; 453 A; 308 C; 363 G; 373 T; 0 other;

Query Match 83.2%; Score 15.8; DB 23; Length 1497;  
Best Local Similarity 89.5%; Pred No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GATCGGCTACCTGACTAC 19  
||| ||||| ||||| |||||  
Db 64 GACCGGCTTCCTGACTAC 46

RESULT 8  
AAF89737  
ID AAF89737 standard; DNA: 2076 BP.  
XX AAF89737;  
AC AAF89737;

XX DT 23-JUL-2001 (first entry)

XX DE Nucleotide sequence of a human metalloprotease enzyme IGS5.

XX KW Metalloprotease; IGS5; infection; pain; cancer; diabetes; obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypertension; hypotension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; stroke; ulcer; allergy; benign prostatic hypertrophy; migraine; psychotic disorder; neurological disorder; autism; multiple sclerosis; Alzheimer's disease; neurodegenerative disease; sleep disorder; epilepsy; kidney disease; cardiovascular disease; arteriosclerosis; cerebrovasospasm; subarachnoid hemorrhage; cerebral ischemia; cerebral infarction; peripheral vascular disease; Raynaud's disease; motility disorder; gastrointestinal disorder; gastric emptying; gastroparesis; diarrhoea; inflammation; chemotherapy induced injury; tumour invasion; immune disorder; arthritis; endotoxin shock; sepsis; diabetes mellitus; severe mental retardation; dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome; ss.

XX OS Homo sapiens.

XX FH Key

XX FT Location/Qualifiers

XX FT 1..2076

XX FT /\*tag= a

XX FT /product= "metalloprotease enzyme IGS5"

XX PN WO200136610-A1.

XX PD 25-MAY-2001.

XX PF 17-NOV-2000; 2000WO-BP11532.

XX PR 19-NOV-1999; 99EP-0203862.

XX PR 19-NOV-1999; 99NL-1013616.

XX PR 31-MAY-2000; 2000EP-0201937.

XX PR 31-MAY-2000; 2000NL-1015356.

XX PA (SOLV ) SOLVAY PHARM BV.

XX PI Deleersnijder W, Wieggers R, Weske M;

XX PX WPI: 2001-343815/36.

XX DR P-PSDB; AAB83840.

XX PT New IGS5 polypeptides useful for treating infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma, schizophrenia, hypertension, urinary retention and Parkinson's disease

XX PS Claim 11; Page 5-6; 115pp; English.

XX CC The present sequence encodes a human metalloprotease enzyme designated IGS5. IGS5 polynucleotides and polypeptides are useful for treating infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, angina pectoris, myocardial infarction, stroke, ulcers, allergies, benign prostatic hypertrophy, migraine, psychotic and neurological disorders, autism, multiple sclerosis, Alzheimer's disease, and other neurodegenerative diseases, sleep disorder, epilepsy, cardiovascular diseases, arteriosclerosis, cerebrovasospasm, subarachnoid hemorrhage, cerebral ischemia, cerebral infarction, peripheral vascular disease, Raynaud's disease, kidney diseases, gastrointestinal disorders, motility disorders and conditions of delayed gastric emptying, post-operative or diabetic gastroparesis, diarrhoea, inflammations, chemotherapy induced injury, tumour invasion, immune disorders, arthritis, endotoxin shock, sepsis, complications of diabetes mellitus, and severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome.

XX SQ Sequence 2076 BP; 493 A; 578 C; 636 G; 369 T; 0 other;

Query Match

83.2%; Score 15.8; DB 22; Length 2076;

Best Local Similarity 89.5%; Pred. No. 1.2e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GATCGGCTACCTGACTAC 19

Db 1257 GATCGGCGACCTGACTAC 1275

||||| ||||| |||||

RESULT 9

AAS97186

ID AAS97186 standard; cDNA; 2232 BP.

XX AC AAS97186;

XX DT 26-FEB-2002 (first entry)

XX DE Human metalloprotease partial DNA sequence #15.

XX KW Human; protease; PCR primer; cytostatic; immunomodulator; cardiant; vasotropic; antimigraine; analgesic; endocrine; nootropic; tranquiliser; hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic; anorectic; antiinflammatory; aspartyl protease; cysteine protease; metalloprotease; serine protease; cancer; haematopoietic; breast; colon; lung; prostate; cervical; brain; ovarian; bladder; kidney; pain; immune-related disease; cardiovascular disease; neuronal disease; migraine; sexual dysfunction; mood disorder; attention disorder; cognition disorder; hypotension; hypertension; psychotic disorder; dyskinesia; metabolic disorder; inflammatory disorder; ss.

XX OS Homo sapiens.

XX PN WO200183782-A2.

XX PD 08-NOV-2001.

XX PF 04-MAY-2001; 2001WO-US14431.

XX PR 04-MAY-2000; 2000US-201879P.

XX PA (SUGE-) SUGEN INC.

XX PI Plozman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;

XX PX Payne V;

XX DR WPI: 2002-041502/05.

XX PT Novel protease polypeptide useful for screening for substances that may be used to treat, e.g., cancers, immune-related diseases, cardiovascular disease, migraine, pain, psychotic and inflammatory disorders

XX PS Claim 30; Figure 1R-S; 232pp; English.

XX CC The invention relates to an isolated, enriched, or purified protease polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to screen for substances (S) that may modulate its activity. Administering S (which modulates protease activity in vitro) may be used to treat a disease or disorder selected from cancers (e.g., of tissues, of blood or haematopoietic origin, of the breast, colon, lung, prostate, cervical, brain, ovarian, bladder or kidney), immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases (e.g., central or peripheral nervous system diseases, migraine, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, hypertension, psychotic disorders, neurological disorders and dyskinesias), metabolic disorders and inflammatory disorders. (I) may also be useful as a diagnostic tool for a disease or disorder such as those above. AAS97159-AAS97195 represent human protease coding sequences and primers of the invention.

XX SQ Sequence 2232 BP; 512 A; 620 C; 705 G; 395 T; 0 other;

Query Match

83.2%; Score 15.8; DB 24; Length 2232;

Best Local Similarity 89.5%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GATCGGCTACCCCTGACTAC 19  
||||| |||||||||

Db 1413 GATCGGGCACCCCTGACTAC 1431

RESULT 10  
AAF89739  
ID AAF89739 standard; DNA; 2262 BP.

XX AAF89739;  
XX  
XX 23-JUL-2001 (first entry)

Nucleotide sequence of a human metalloprotease enzyme IGS5.

XX Metalloprotease; IGS5; infection; pain; cancer; diabetes; obesity;  
KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;  
KW hypotension; hypertension; urinary retention; osteoporosis;  
KW angina pectoris; myocardial infarction; stroke; ulcer; allergy;  
KW benign prostatic hypertrophy; migraine; psychotic disorder;  
KW neurological disorder; autism; multiple sclerosis; Alzheimer's disease;  
KW neurodegenerative disease; sleep disorder; epilepsy; kidney disease;  
KW cardiovascular disease; arteriosclerosis; cerebrovasospasm;  
KW subarachnoid hemorrhage; cerebral ischemia; cerebral infarction;  
KW peripheral vascular disease; Raynaud's disease; motility disorder;  
KW gastrointestinal disorder; gastric emptying; gastroparesis; diarrhoea;  
KW inflammation; chemotherapy induced injury; tumour invasion;  
KW immune disorder; arthritis; endotoxin shock; sepsis; diabetes mellitus;  
KW severe mental retardation; dyskinesia; Huntington's disease;  
KW Gilles de la Tourette's syndrome; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
FT CDS 1..2262  
FT /tag= a  
FT /product= "metalloprotease enzyme IGS5"

XX WO200136610-A1.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-EF11532.

XX 19-NOV-1999; 99EP-0203862.  
XX 19-NOV-1999; 99NL-1013616.  
XX 31-MAY-2000; 2000EP-0201937.  
XX 31-MAY-2000; 2000NL-1015356.

XX (SOLV ) SOLVAY PHARM BV.

XX Deleersnijder W, Wiegiers R, Weske M;  
XX  
XX WPI; 2001-343815/36.  
XX P-PSDB; AAB83842.

XX New IGS5 polypeptides useful for treating infections, pain, cancer,  
XX diabetes, obesity, anorexia, bulimia, asthma, schizophrenia,  
XX hypertension, urinary retention and Parkinson's disease  
XX  
XX Claim 11; Page 8-9; 115pp; English.

XX The present sequence encodes a human metalloprotease enzyme designated.  
XX IGS5. IGS5 polynucleotides and polypeptides are useful for treating  
XX infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma,  
XX Parkinson's disease, acute heart failure, hypotension, hypertension,  
XX urinary retention, osteoporosis, angina pectoris, myocardial infarction,  
XX stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,  
XX psychotic and neurological disorders, autism, multiple sclerosis,  
XX Alzheimer's disease, and other neurodegenerative diseases, sleep

CC disorder, epilepsy, cardiovascular diseases, arteriosclerosis,  
CC cerebrovasospasm, subarachnoid hemorrhage, cerebral ischemia, cerebral  
CC infarction, peripheral vascular disease, Raynaud's disease, kidney  
CC diseases, gastrointestinal disorders, motility disorders and conditions  
CC of delayed gastric emptying, post-operative or diabetic gastroparesis,  
CC diarrhoea, inflammations, chemotherapy induced injury, tumour invasion,  
CC immune disorders, arthritis, endotoxin shock, sepsis, complications of  
CC diabetes mellitus, and severe mental retardation and dyskinesias, such  
CC as Huntington's disease or Gilles de la Tourette's syndrome.

XX  
XX SQ Sequence 2262 BP; 520 A; 628 C; 716 G; 398 T; 0 other;

Query Match 83.2%; Score 15.8; DB 22;  
Best Local Similarity 89.5%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GATCGGCTACCCCTGACTAC 19  
||||| |||||||||

Db 1443 GATCGGGCACCCCTGACTAC 1461

RESULT 11  
AAD30580  
ID AAD30580 standard; cDNA; 2318 BP.

XX AAD30580;  
XX  
XX 21-MAY-2002 (first entry)

XX Human protease, PRYS-13 cDNA.

XX Human; protease; PRYS-13; enzyme; gastritis; cirrhosis; Crohn's disease;  
KW gastrointestinal disorder; autoimmune; inflammatory; cell proliferative;  
KW cardiovascular; developmental; epithelial; neurological; reproductive;  
KW AIDS; Acquired Immune Deficiency Syndrome; allergy; rheumatoid arthritis;  
KW anaemia; asthma; arteriosclerosis; hypertension; myocardial infarction;  
KW hepatitis; cancer; psoriasis; Cushing's syndrome; hypothyroidism; eczema;  
KW epilepsy; Alzheimer's disease; Huntington's disease; Parkinson's disease;  
KW Pick's disease; infertility; vitiligo; drug screening; gene therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
FT CDS 338..1651  
FT /tag= a  
FT /product= "Human PRYS-13 protein"

XX sig\_peptide 338..427  
FT /tag= b  
FT mat\_peptide 428..1648  
FT /tag= c  
FT /product= "Mature PRYS-13 protein"

XX WO200208396-A2.

XX 31-JAN-2002.

XX 17-JUL-2001; 2001WO-US22397.

XX 21-JUL-2000; 2000US-220063P.  
XX 28-JUL-2000; 2000US-221680P.  
XX 04-AUG-2000; 2000US-223544P.  
XX 11-AUG-2000; 2000US-224717P.  
XX 16-AUG-2000; 2000US-225988P.  
XX 23-AUG-2000; 2000US-227568P.

XX (INCY-) INCYTE GENOMICS INC.

XX Deleage AM, Gandhi AR, Hafalia AJA, Lu DAM, Patterson C;  
XX Tribouley CM, Das D, Kallick DA, Nguyen DB, Lee EA, Khan FA;  
XX Yue H, Au-Young J, Griffin JA, Policky JL, Ramkumar J, Yang J;  
XX Thangavelu K, Ding L, Kearney L, Baughn MR, Borowsky ML;  
XX Sanjanwala MS, Yao MG, Burford N, Walla NK, Lal P, Lee S, Todd S;  
XX Lo TP, Tang YT, Elliott VS, Azimzai Y, Lu Y;

XX WPI: 2002-206082/26.  
 DR P-PSDB; AAE19176.  
 XX  
 PT New human protease polypeptide, useful in diagnosis, prevention and  
 PT treatment of gastrointestinal, cardiovascular, autoimmune/inflammatory,  
 PT cell proliferative, developmental, epithelial and neurological  
 PT disorders  
 XX  
 PS Claim 5; Page 174-175; 182pp; English.  
 XX  
 XX The invention relates to an isolated human protease polypeptide (PRTS).  
 CC PRTS protein and DNA are useful for diagnosing, treating and preventing  
 CC gastrointestinal disorders (gastritis, cirrhosis, Crohn's disease),  
 CC autoimmune/inflammatory disorders (AIDS, allergy, rheumatoid arthritis,  
 CC anaemia, asthma), cardiovascular disorder (atherosclerosis, hypertension,  
 CC myocardial infarction), cell proliferative disorders (hepatitis, cancer,  
 CC psoriasis), developmental disorders (Cushing's syndrome, hypothyroidism,  
 CC epithelial disorder (vitiligo, keloid, eczema), neurological disorders  
 CC (epilepsy, Alzheimer's disease, Pick's disease, Huntington's disease,  
 CC Parkinson's disease), and reproductive disorders (infertility). PRTS  
 CC protein is useful in a number of drug screening techniques and to  
 CC analyse the proteome of a tissue or cell type. PRTS DNA is useful for  
 CC creating knockin humanised animals or transgenic animals to model human  
 CC diseases, in somatic or germline gene therapy and in microarrays  
 CC utilising fluids or tissues from patients to detect altered PRIN  
 CC expression. The present sequence is human PRTS-13 cDNA.  
 XX  
 SQ Sequence 2318 BP; 526 A; 647 C; 735 G; 410 T; 0 other;  
 Query Match 83.2%; Score 15.8; DB 24; Length 2318;  
 Best Local Similarity 89.5%; Pred. No. 1.2e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GATCGGCTACCTGACTAC 19  
 ||||| |||||  
 DB 1502 GATCGGCACCTGACTAC 1520  
 RESULT 12  
 AAF89738  
 ID AAF89738 standard; DNA; 2340 BP.  
 XX  
 AC AAF89738;  
 XX  
 DT 23-JUL-2001 (first entry)  
 XX  
 DE Nucleotide sequence of a human metalloprotease enzyme IGS5.  
 KW Metalloprotease; IGS5; infection; pain; cancer; diabetes; obesity;  
 KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;  
 KW hypotension; hypertension; urinary retention; osteoporosis;  
 KW angina pectoris; myocardial infarction; stroke; ulcer; allergy;  
 KW benign prostatic hypertrophy; migraine; psychotic disorder;  
 KW neurological disorder; autism; multiple sclerosis; Alzheimer's disease;  
 KW neurodegenerative disease; sleep disorder; epilepsy; kidney disease;  
 KW cardiovascular disease; arteriosclerosis; cerebrovasospasm;  
 KW subarachnoid hemorrhage; cerebral ischemia; cerebral infarction;  
 KW peripheral vascular disease; Raynaud's disease; motility disorder;  
 KW gastrointestinal disorder; gastric emptying; gastroparesis; diarrhoea;  
 KW inflammation; chemotherapy induced injury; tumour invasion;  
 KW immune disorder; arthritis; endotoxin shock; sepsis; diabetes mellitus;  
 KW severe mental retardation; dyskinesia; Huntington's disease;  
 KW Gilles de la Tourette's syndrome; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key. Location/Qualifiers  
 FT CDS 1..2340  
 FT /\*tag= a  
 FT /product= "metalloprotease enzyme IGS5"  
 FT  
 PN WO200136610-A1.

XX 25-MAY-2001.  
 XX PD  
 XX PF  
 XX 17-NOV-2000; 2000WO-EPI1532.  
 XX  
 PR 19-NOV-1999; 99EP-0203862.  
 PR 19-NOV-1999; 99NL-1013616.  
 PR 31-MAY-2000; 2000EP-0201937.  
 PR 31-MAY-2000; 2000NL-1015356.  
 XX  
 PA (SOLV ) SOLVAY PHARM BV.  
 XX  
 PI Deleersnijder W, Wiegers R, Weske M;  
 XX  
 DR WPI: 2001-343815/36.  
 DR P-PSDB; AAB83841.  
 XX  
 PT New IGS5 polypeptides useful for treating infections, pain, cancer,  
 PT diabetes, obesity, anorexia, bulimia, asthma, schizophrenia,  
 PT hypertension, urinary retention and Parkinson's disease  
 XX  
 PS Claim 11; Page 6-7; 115pp; English.  
 XX  
 XX The present sequence encodes a human metalloprotease enzyme designated  
 CC IGS5. IGS5 polynucleotides and polypeptides are useful for treating  
 CC infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma,  
 CC Parkinson's disease, acute heart failure, hypotension, hypertension,  
 CC urinary retention, osteoporosis, angina pectoris, myocardial infarction,  
 CC stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,  
 CC psychotic and neurological disorders, autism, multiple sclerosis,  
 CC Alzheimer's disease, and other neurodegenerative diseases, sleep  
 CC disorder, epilepsy, cardiovascular diseases, arteriosclerosis,  
 CC cerebrovasospasm, subarachnoid hemorrhage, cerebral ischemia, cerebral  
 CC infarction, peripheral vascular disease, Raynaud's disease, kidney  
 CC diseases, gastrointestinal disorders, motility disorders and conditions  
 CC of delayed gastric emptying, post-operative or diabetic gastroparesis,  
 CC diarrhoea, inflammations, chemotherapy induced injury, tumour invasion,  
 CC immune disorders, arthritis, endotoxin shock, sepsis, complications of  
 CC diabetes mellitus, and severe mental retardation and dyskinesias, such  
 CC as Huntington's disease or Gilles de la Tourette's syndrome.  
 XX  
 SQ Sequence 2340 BP; 539 A; 649 C; 739 G; 413 T; 0 other;  
 Query Match 83.2%; Score 15.8; DB 22; Length 2340;  
 Best Local Similarity 89.5%; Pred. No. 1.2e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GATCGGCTACCTGACTAC 19  
 ||||| |||||  
 DB 1521 GATCGGCACCTGACTAC 1539  
 RESULT 13  
 ABL22299/c  
 ID ABL22299 standard; DNA; 2529 BP.  
 XX  
 AC ABL22299;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 18370.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ds.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX

PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX Claim 1; SEQ ID NO 18370; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
CC sequences (AB101840-AB116175) and the encoded proteins  
CC (AB157737-AB172072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 2529 BP; 752 A; 721 C; 566 G; 490 T; 0 other;  
SQ Query Match 83.2%; Score 15.8; DB 23; Length 2529;  
Best Local Similarity 89.5%; Pred. No. 1.3e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 GATCGGTACCTGACTAC 19  
Db 1663 GATCGGTACCTGACTAC 1645  
RESULT 14  
ABN84280  
ID ABN84280 standard; cDNA; 2580 BP.  
XX  
XX AC ABN84280;  
XX  
XX 23-SEP-2002 (first entry)  
XX Human SEP endopeptidase coding sequence.  
XX SEP; endopeptidase; enzyme; human; neuropeptide Y1; receptor;  
KW male sexual dysfunction; male erectile dysfunction; obesity;  
KW anorexia; bulimia; vasotropic; anorectic; therapy; inhibitor; gene;  
KW ss.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH misc\_feature 1..65  
FT /\*tag= a  
FT /\*note= "5' partial vector sequence"  
FT CDS 258..2580  
FT /\*tag= b  
FT /\*product= "SEP"  
FT /\*partial  
FT /\*note= "the CDS does not include a stop codon"  
FT 2580  
FT /\*tag= c  
FT /\*note= "the 3' end of the sequence is missing  
FT from Figure 8 of the specification"  
XX  
XX WO200247670-A1.  
XX  
XX 20-JUN-2002.  
XX

PF 10-DEC-2001; 2001WO-IB02399.  
XX  
XX 15-DEC-2000; 2000GB-0030647.  
XX 06-APR-2001; 2001GB-0008730.  
XX 23-APR-2001; 2001GB-0009910.  
XX 04-MAY-2001; 2001GB-0011037.  
XX 29-JUN-2001; 2001US-0895367.  
XX 13-JUL-2001; 2001US-0905846.  
XX 24-AUG-2001; 2001GB-0020679.  
XX (PFIZ ) PFIZER LTD.  
XX (PFIZ ) PFIZER INC.  
XX  
XX Naylor AM, Van Der Graaf PH, Wayman CP;  
XX  
XX WPI; 2002-547828/58.  
XX P-PSDB; ABB79521.  
XX  
XX Use of an inhibitor of neuropeptide Y in the preparation of medicament  
PT for the treatment or prevention of male erectile dysfunction -  
XX  
XX Disclosure; Fig 8; 179pp; English.  
XX  
XX The present sequence is a nucleotide sequence coding for human  
CC SEP, a soluble secreted endopeptidase, with an additional 5'  
CC vector partial sequence. The invention relates to the use of an  
CC inhibitor of NPY, or an inhibitor of NPY Y1 receptor, especially an  
CC inhibitor selective for an NPY or NPY Y1 receptor associated with  
CC male genitalia, in the preparation of a medicament for the  
CC treatment or prevention of male sexual dysfunction, especially male  
CC erectile dysfunction (MED). The NPY inhibitor may be used with an  
CC auxiliary active agent such as an SEP inhibitor. The invention  
CC provides a SEP assay that can be used to detect candidate  
CC inhibitors of SEP. In addition to treatment of MED, NPY inhibitors  
CC can also be used to treat abnormal drink and food intake disorders,  
CC such as obesity, bulimia, anorexia and metabolic disorders.  
XX  
XX Sequence 2580 BP; 587 A; 741 C; 794 G; 458 T; 0 other;  
SQ Query Match 83.2%; Score 15.8; DB 24; Length 2580;  
Best Local Similarity 89.5%; Pred. No. 1.3e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 GATCGGTACCTGACTAC 19  
Db 1770 GATCGGTACCTGACTAC 1788  
RESULT 15  
AAF59660  
ID AAF59660 standard; cDNA; 2636 BP.  
XX  
XX AC AAF59660;  
XX  
XX 27-APR-2001 (first entry)  
XX  
XX Human neprilysin-like membrane metalloproteinase SNEPb cDNA.  
XX  
XX Human; SNEPb; neprilysin-like membrane metalloproteinase;  
KW splice variant; alternative splicing; zinc endopeptidase family;  
KW neuropeptide; peptide hormone; processing; metabolism; vaccine; pain;  
KW drug screening; renal insufficiency; hepatic disease; ischaemia; pain;  
KW hypertension; cancer; inflammation; cardiovascular disease;  
KW neuronal disease; pancreatic disease; prostatic disease;  
KW respiratory disease; gene therapy; nephrotropic; vasotropic; analgesic;  
KW cerebroprotective; hypotensive; cytostatic; antiinflammatory;  
KW cardiovascular; hepatotropic; ss.  
XX  
XX Homo sapiens.  
XX  
XX EP1069188-A1.  
XX  
XX 17-JAN-2001.  
XX

XX 15-JUL-1999; 99EP-0401767.  
 XX 15-JUL-1999; 99EP-0401767.  
 XX (SNFI ) SANOFI-SYNTHELABO.  
 XX Jagerschmidt A, Agnel M, Culouscou J;  
 XX WPI; 2001-212582/22.  
 XX P-PSDB; AAB60562.  
 XX New membrane-associated metalloproteinase SNEPa, SNEPb and SNEPc  
 PT polypeptides and polynucleotides, useful for treating e.g. acute and  
 PT chronic renal insufficiency, pain, stroke, cancer, inflammation, renal  
 PT and hepatic ischemia  
 XX  
 PS Claim 5; Page 30-33; 72pp; English.  
 XX  
 CC The invention relates to the human SNEPa, SNEPb and SNEPc proteins, and  
 CC the cDNAs encoding them. SNEPa, SNEPb and SNEPc are neprilysin-like  
 CC membrane metalloproteinases and are the products of alternative splicing.  
 CC The substrate(s) for the SNEP proteins are not as yet known, although  
 CC the neprilysin family of zinc endopeptidases play key roles in the  
 CC processing and/or metabolism of neuropeptides and peptide hormones. SNEP  
 CC nucleotides may be used as hybridisation probes for cDNA and genomic  
 CC DNA; to isolate full-length cDNAs and genomic clones encoding SNEPa,  
 CC SNEPb or SNEPc; to isolate cDNA and genomic clones of SNEP homologues;  
 CC as research reagents and material for the discovery of treatments and  
 CC diagnostics for animal and human diseases; and for chromosome  
 CC identification. The SNEP proteins may be used as immunogens to  
 CC produce antibodies immunospecific for SNEPa, SNEPb or SNEPc. Such  
 CC antibodies are used to isolate or identify clones expressing the  
 CC protein, or to purify the proteins by affinity chromatography.  
 CC SNEP proteins may also be used in screening for compounds which modulate  
 CC SNEP endopeptidase activity, and to assess enzymatic cleavage of small  
 CC molecule substrates in cells, cell-free preparations, chemical libraries  
 CC and product mixtures. The SNEP proteins (as vaccine compositions),  
 CC SNEP nucleotides, and SNEP activators or inhibitors may be used  
 CC to treat acute and chronic renal insufficiency, renal and hepatic  
 CC ischaemia, pain, stroke, hypertensive disease, cancer, inflammation, as  
 CC well as cardiovascular, neuronal, pancreatic, prostatic, renal,  
 CC respiratory or hepatic diseases. They may also be used in modulating  
 CC peptide activation and/or degradation in the brain or kidney or in  
 CC another organ, or to diagnose or treat any disorder related to abnormal  
 CC expression of SNEPa, SNEPb or SNEPc. The present sequence represents  
 CC cDNA encoding SNEPb.  
 XX  
 SQ Sequence 2636 BP; 601 A; 759 C; 813 G; 462 T; 1 other;  
 Query Match 83.2%; Score 15.8; DB 22; Length 2636;  
 Best Local Similarity 89.5%; Pred. No. 1.3e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GATCGGCTACCTGACTAC 19  
 Db 1450 GATCGGCGACCTGACTAC 1468  
 ||||| - |||||

Search completed: July 8, 2003, 02:19:13  
 Job time : 120.376 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 01:24:03 ; Search time 26.2293 seconds

(without alignments)  
222.151 Million cell updates/sec

Title: US-09-647-780A-19

Perfect score: 19

Sequence: 1 gatcggtaccctgactac 19

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA: \*  
1: /cn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cn2\_6/ptodata/1/ina/PTUS\_COMB.seq: \*  
6: /cn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	14.4	75.8	4403765	4	US-09-103-840A-2
c 2	14.4	75.8	4411529	4	US-09-103-840A-1
c 3	14.2	74.7	567	1	US-08-463-115-6
4	14.2	74.7	567	1	US-08-465-388-6
5	14.2	74.7	606	4	US-09-305-640-3
6	14.2	74.7	666	1	US-08-463-115-4
c 7	14.2	74.7	666	1	US-08-463-115-4
8	14.2	74.7	666	1	US-08-465-388-4
c 9	14.2	74.7	666	1	US-08-465-388-4
10	14.2	74.7	2560	4	US-09-305-640-1
c 11	14.2	74.7	4224	1	US-08-612-521-1
12	14.2	74.7	5057	2	US-08-365-486A-12
13	14.2	74.7	5057	4	US-08-880-342-12
14	14.2	74.7	5108	1	US-07-642-002-1
15	14.2	74.7	5859	1	US-08-312-387B-1
16	14.2	74.7	5859	1	US-08-312-387B-7
17	14.2	74.7	5859	1	US-08-683-426-1
18	14.2	74.7	5859	1	US-08-683-426-7
19	14.2	74.7	5859	1	US-08-683-458-1
20	14.2	74.7	5859	1	US-08-683-458-7
21	14.2	74.7	5859	2	US-08-878-360-1
22	14.2	74.7	5859	2	US-08-878-360-7
23	14.2	74.7	5859	3	US-08-478-140B-1
24	14.2	74.7	5859	4	US-09-333-412-1
25	14.2	74.7	5859	4	US-09-333-412-7
26	14.2	74.7	5859	4	US-09-338-943-1
c 27	14.2	74.7	6143	1	US-08-612-521-3

28	14.2	74.7	15664	1	US-08-402-282-3	Sequence 3, Appli
29	14.2	74.7	15664	1	US-08-508-004-3	Sequence 3, Appli
30	14.2	74.7	15664	1	US-08-402-066-3	Sequence 3, Appli
31	14.2	74.7	15664	1	US-08-402-068-3	Sequence 3, Appli
c 32	14.2	74.7	80161	3	US-09-036-987A-1	Sequence 1, Appli
c 33	14.2	74.7	80161	4	US-09-370-700-1	Sequence 1, Appli
34	14	73.7	4247	1	US-08-061-465-2	Sequence 2, Appli
35	14	73.7	5828	1	US-08-061-465-1	Sequence 1, Appli
36	13.8	72.6	491	3	US-08-361-441B-43	Sequence 43, Appli
37	13.8	72.6	606	4	US-09-280-116-121	Sequence 121, App
38	13.8	72.6	658	4	US-08-792-013-5	Sequence 5, Appli
39	13.8	72.6	672	1	US-08-485-455D-16	Sequence 16, Appl
40	13.8	72.6	672	2	US-08-482-130C-16	Sequence 16, Appl
41	13.8	72.6	672	2	US-08-484-211C-16	Sequence 16, Appl
42	13.8	72.6	672	3	US-08-906-769-16	Sequence 16, Appl
43	13.8	72.6	672	3	US-08-906-616-16	Sequence 16, Appl
44	13.8	72.6	672	3	US-08-817-795-16	Sequence 16, Appl
45	13.8	72.6	672	3	US-08-485-443B-16	Sequence 16, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328

; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103.840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match 75.8%; Score 14.4; DB 4; Length 4403765;  
Best Local Similarity 93.8%; Pred. No. 66;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CGGCTACCTGACTAC 19  
|||||||

Db 478219 CGGCTACCTGACTAC 478204

##### RESULT 2

US-09-103-840A-1/c  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328

; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103.840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 4411529  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
OTHER INFORMATION: H37RV  
US-09-103-840A-1

Query Match 75.8%; Score 14.4; DB 4; Length 4411529;  
Best Local Similarity 93.8%; Pred. No. 66;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CGGCTACCCCTGACTAC 19  
|||||  
Db 478135 CGGCTACCCGCTACT 478120

RESULT 3  
US-08-463-115-6  
; Sequence 6, Application US/08463115  
; Patent No. 5703221  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAM JOHN MARTIN  
; TITLE OF INVENTION: ISOLATED STEALTH VIRUSES  
; TITLE OF INVENTION: AND RELATED VACCINES  
; NUMBER OF SEQUENCES: 104  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,115  
; FILING DATE: June 5, 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; PRIOR APPLICATION DATA: including application  
; PRIOR APPLICATION DATA: described below:  
; APPLICATION NUMBER: 08/157,811  
; FILING DATE: No. 5703221ember 23, 1993  
; APPLICATION NUMBER: 07/887,502  
; FILING DATE: May 22, 1992  
; APPLICATION NUMBER: 07/704,814  
; FILING DATE: May 23, 1991  
; APPLICATION NUMBER: 07/763,039  
; FILING DATE: September 20, 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 213/301  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 567 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-463-115-6

Query Match 74.7%; Score 14.2; DB 1; Length 567;  
Best Local Similarity 84.2%; Pred. No. 1.2e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 GATCGGCTACCCCTGACTAC 19  
|||||  
Db 7 GATCGGATACCCCGTCTAC 25

RESULT 4  
US-08-465-388-6  
; Sequence 6, Application US/08465388  
; Patent No. 5753488  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAM JOHN MARTIN  
; TITLE OF INVENTION: ISOLATED STEALTH VIRUSES  
; TITLE OF INVENTION: AND RELATED VACCINES  
; NUMBER OF SEQUENCES: 104  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465,388  
; FILING DATE: June 5, 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; PRIOR APPLICATION DATA: including application  
; PRIOR APPLICATION DATA: described below:  
; APPLICATION NUMBER: 08/157,811  
; FILING DATE: No. 5753488ember 23, 1993  
; APPLICATION NUMBER: 07/887,502  
; FILING DATE: May 22, 1992  
; APPLICATION NUMBER: 07/704,814  
; FILING DATE: May 23, 1991  
; APPLICATION NUMBER: 07/763,039  
; FILING DATE: September 20, 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 213/300  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 567 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-465-388-6

Query Match 74.7%; Score 14.2; DB 1; Length 567;  
Best Local Similarity 84.2%; Pred. No. 1.2e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GATCGGCTACCCCTGACTAC 19  
|||||  
Db 7 GATCGGATACCCCGTCTAC 25

RESULT 5  
US-09-305-640-3  
; Sequence 3, Application US/09305640B



Patent No. 6255468  
GENERAL INFORMATION:  
APPLICANT: SmithKline Beecham plc  
TITLE OF INVENTION: No. 6255468el Compounds  
FILE REFERENCE: GP30124  
CURRENT APPLICATION NUMBER: US/09/305,640B  
CURRENT FILING DATE: 1999-05-05  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 606  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (404)(437)(460)(490)(520)(555)(561)(595)  
US-09-305-640-3

Query Match 74.7%; Score 14.2; DB 4; Length 606;  
Best Local Similarity 84.2%; Pred. No. 1.2e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATCGGTACCTCGACTAC 19  
|||||  
DB 1 GGTGGGTACCCGACTTC 19

RESULT 6  
US-08-463-115-4  
Sequence 4, Application US/08463115  
Patent No. 5703221  
GENERAL INFORMATION:  
APPLICANT: WILLIAM JOHN MARTIN  
TITLE OF INVENTION: ISOLATED STEALTH VIRUSES  
TITLE OF INVENTION: AND RELATED VACCINES  
NUMBER OF SEQUENCES: 104  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,115  
FILING DATE: June 5, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: 08/157,811  
FILING DATE: No. 5703221ember 23, 1993  
APPLICATION NUMBER: 07/887,502  
FILING DATE: May 22, 1992  
APPLICATION NUMBER: 07/704,814  
FILING DATE: May 23, 1991  
FILING DATE: September 20, 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 213/301  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 666 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-463-115-4

Query Match 74.7%; Score 14.2; DB 1; Length 666;  
Best Local Similarity 84.2%; Pred. No. 1.2e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATCGGTACCTCGACTAC 19  
|||||  
DB 8 GATCGGTACCCGCTCTAC 26

RESULT 7  
US-08-463-115-4/c  
Sequence 4, Application US/08463115  
Patent No. 5703221  
GENERAL INFORMATION:  
APPLICANT: WILLIAM JOHN MARTIN  
TITLE OF INVENTION: ISOLATED STEALTH VIRUSES  
TITLE OF INVENTION: AND RELATED VACCINES  
NUMBER OF SEQUENCES: 104  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,115  
FILING DATE: June 5, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: 08/157,811  
FILING DATE: No. 5703221ember 23, 1993  
APPLICATION NUMBER: 07/887,502  
FILING DATE: May 22, 1992  
APPLICATION NUMBER: 07/704,814  
FILING DATE: May 23, 1991  
FILING DATE: September 20, 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 213/301  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 666 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-463-115-4

Query Match 74.7%; Score 14.2; DB 1; Length 666;  
Best Local Similarity 84.2%; Pred. No. 1.2e+02;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GATCGGTACCTGACTAC 19  
||||| ||||| |||||  
Db 660 GATCGGATACCCGCTCTAC 642

## RESULT 8

US-08-465-388-4  
; Sequence 4, Application US/08465388  
; Patent No. 5753488  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAM JOHN MARTIN  
; TITLE OF INVENTION: ISOLATED STEALTH VIRUSES  
; NUMBER OF SEQUENCES: 104  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465,388  
; FILING DATE: June 5, 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA: including application  
; PRIOR APPLICATION DATA: described below:  
; APPLICATION NUMBER: 08/157,811  
; FILING DATE: No. 5753488ember 23, 1993  
; APPLICATION NUMBER: 07/887,502  
; FILING DATE: May 22, 1992  
; APPLICATION NUMBER: 07/704,814  
; FILING DATE: May 23, 1991  
; APPLICATION NUMBER: 07/763,039  
; FILING DATE: September 20, 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 213/300  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 666 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-465-388-4

Query Match 74.7%; Score 14.2; DB 1; Length 666;  
Best Local Similarity 84.2%; Pred. No. 1.2e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GATCGGTACCTGACTAC 19  
||||| ||||| |||||  
Db 8 GATCGGATACCCGCTCTAC 26

## RESULT 9

US-08-465-388-4/c  
; Sequence 4, Application US/08465388

; Patent No. 5753488  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAM JOHN MARTIN  
; TITLE OF INVENTION: ISOLATED STEALTH VIRUSES  
; NUMBER OF SEQUENCES: 104  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465,388  
; FILING DATE: June 5, 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA: including application  
; PRIOR APPLICATION DATA: described below:  
; APPLICATION NUMBER: 08/157,811  
; FILING DATE: No. 5753488ember 23, 1993  
; APPLICATION NUMBER: 07/887,502  
; FILING DATE: May 22, 1992  
; APPLICATION NUMBER: 07/704,814  
; FILING DATE: May 23, 1991  
; APPLICATION NUMBER: 07/763,039  
; FILING DATE: September 20, 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 213/300  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 666 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-465-388-4

Query Match 74.7%; Score 14.2; DB 1; Length 666;  
Best Local Similarity 84.2%; Pred. No. 1.2e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GATCGGTACCTGACTAC 19  
||||| ||||| |||||  
Db 660 GATCGGATACCCGCTCTAC 642

## RESULT 10

US-09-305-640-1  
; Sequence 1, Application US/09305640B  
; Patent No. 6255468  
; GENERAL INFORMATION:  
; APPLICANT: SmithKline Beecham plc  
; TITLE OF INVENTION: No. 6255468el Compounds  
; FILE REFERENCE: GP30124  
; CURRENT APPLICATION NUMBER: US/09/305,640B  
; CURRENT FILING DATE: 1999-05-05  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1

```

; LENGTH: 2560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-305-640-1

Query Match 74.7%; Score 14.2; DB 4; Length 2560;
Best Local Similarity 84.2%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATCGGCTACCCGTGACTAC 19
| | | | | | | | | | | | | | | | | | | | | |
Db 1752 GTCGCGTACCCGCGACTTC 1770

RESULT 11
US-08-612-521-1/C
; Sequence 1, Application US/08612521-
; Patent No. 5786463
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B
; APPLICANT: Skatrud, Paul L
; APPLICANT: Thorneswell, Susan J
; TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF
; TITLE OF INVENTION: CRYPTOCOCCUS NEOFORMANS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patent Division/AEH
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,521
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamilton, Amy E
; REGISTRATION NUMBER: 33,894
; REFERENCE/DOCKET NUMBER: X-9693
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3169
; TELEFAX: 317-276-1294
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4224 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4224
US-08-612-521-1

Query Match 74.7%; Score 14.2; DB 1; Length 4224;
Best Local Similarity 84.2%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATCGGCTACCCGTGACTAC 19
| | | | | | | | | | | | | | | | | | | | | |
Db 4147 GTCGCGTACCCGTGCTTC 4129

RESULT 12
US-08-365-486A-12
; Sequence 12, Application US/08365486A
; Patent No. 5834306
; GENERAL INFORMATION:
; APPLICANT: Webster, Keith A.
; APPLICANT: Bishopric, Nanette H.
; APPLICANT: Murphy, Brian
; APPLICANT: Laderoute, Keith R.
; APPLICANT: Green, Christopher J.
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
; TITLE OF INVENTION: Therapeutic Constructs
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306

```

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/880.342  
FILING DATE: 23-JUN-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/IB95/00996  
FILING DATE: 13-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/365,486  
FILING DATE: 23-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 8255-0018.30  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5057 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cdna to mrna  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: rat bnos cdna  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 349..4638  
US-08-880-342-12

Query Match 74.7%; Score 14.2; DB 4; Length 5057;  
Best Local Similarity 84.2%; Pred. No. 1.5e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATCGGCTACCTGACTAC 19  
||||||| || |||||  
DB 2124 GATCGGCTCGTGACTAC 2142

RESULT 14  
US-07-642-002-1  
Sequence 1, Application US/07642002  
Patent No. 5268465  
GENERAL INFORMATION:  
APPLICANT: Bredt, David S.  
APPLICANT: Hwang, Paul M.  
APPLICANT: Reed, Randall  
APPLICANT: Snyder, Solomon H.  
TITLE OF INVENTION: Purification and Molecular Cloning of Nitric  
TITLE OF INVENTION: Oxide Synthase  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie & Beckett  
STREET: One Thomas Circle, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/642,002

FILING DATE: 19910118  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.033576  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 296-5500  
TELEFAX: (202) 296-7830  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5108 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
HYPOTHETICAL: N  
ANTI-SENSE: N  
ORIGINAL SOURCE:  
ORGANISM: Rattus rattus  
TISSUE TYPE: Brain  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 400..4686  
OTHER INFORMATION:  
US-07-642-002-1  
Query Match 74.7%; Score 14.2; DB 1; Length 5108;  
Best Local Similarity 84.2%; Pred. No. 1.5e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATCGGCTACCTGACTAC 19  
||||||| || |||||  
DB 2175 GATCGGCTCGTGACTAC 2193

RESULT 15  
US-08-312-387B-1  
Sequence 1, Application US/08312387B  
Patent No. 5545553  
GENERAL INFORMATION:  
APPLICANT: Gotschlich, Emil C.  
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF  
TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/312,387B  
FILING DATE: July 7, 1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-095  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5859 base pairs  
TYPE: nucleic acid

;; STRANDEDNESS: both  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: DNA (genomic)  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; ORIGINAL SOURCE:  
;; ORGANISM: Neisseria gonorrhoeae.  
;; STRAIN: F62  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..381  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 445..1491  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 2342..3262  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 3322..4335  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 4354..5196  
;;  
US-08-312-387B-1

Query Match 74.7%; Score 14.2; DB 1; Length 5859;  
Best Local Similarity 84.2%; Pred. No. 1.5e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GATCGGCTACCCCTGACTAC 19  
| | | | | | | | | | | | | | | | | | | | | |  
Db 2045 GTTCGGCAACCCCTGACGAC 2063

Search completed: July 8, 2003, 09:32:41  
Job time : 38.2793 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 19:09:54 ; Search time 109.737 seconds  
(without alignments)  
273.390 Million cell updates/sec

Title: US-09-647-780A-19  
Perfect score: 19  
Sequence: 1 gatcggtacctgactac 19

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1105431 seqs, 789497651 residues  
Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_applications\_NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US05\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15.8	83.2	2893	9 US-10-017-273A-4	Sequence 4, Appli
2	15.8	83.2	2893	10 US-09-905-846-1	Sequence 1, Appli
3	15.8	83.2	2975	9 US-10-017-273A-5	Sequence 5, Appli
4	15.8	83.2	2975	10 US-09-905-846-5	Sequence 5, Appli
5	15.4	81.1	432	9 US-09-938-842A-2328	Sequence 2328, Ap
6	15.4	81.1	574	9 US-09-910-664-65	Sequence 65, Appli
7	15.4	81.1	967	10 US-09-770-445-294	Sequence 294, App
8	14.8	77.9	25	9 US-10-098-263B-92943	Sequence 92943, A
9	14.8	77.9	306	10 US-09-294-093B-4481	Sequence 4481, Ap
10	14.8	77.9	5250	9 US-10-091-483-326	Sequence 326, App
11	14.8	77.9	5250	9 US-09-764-891-8224	Sequence 8224, Ap
12	14.8	77.9	5250	10 US-09-764-846-326	Sequence 326, App
13	14.4	75.8	25	9 US-10-098-263B-97871	Sequence 97871, A
14	14.4	75.8	840	9 US-10-156-761-4455	Sequence 4455, Ap
15	14.4	75.8	4053	9 US-10-037-270-156	Sequence 156, App
16	14.4	75.8	9025608	9 US-10-156-761-1	Sequence 1, Appli
17	14.2	74.7	25	9 US-10-098-263B-51979	Sequence 51979, A
18	14.2	74.7	200	9 US-10-046-935-510	Sequence 510, App
19	14.2	74.7	200	9 US-09-878-178-510	Sequence 510, App

Sequence 510, App  
Sequence 550, App  
Sequence 1019, Ap  
Sequence 1019, Ap  
Sequence 1019, Ap  
Sequence 1019, Ap  
Sequence 17948, A  
Sequence 3748, Ap  
Sequence 37923, A  
Sequence 37957, A  
Sequence 9905, Ap  
Sequence 67, Appli  
Sequence 67, Appli  
Sequence 99, Appli  
Sequence 101, App  
Sequence 107, App  
Sequence 109, App  
Sequence 523, App  
Sequence 1151, Ap  
Sequence 124, App  
Sequence 122, App  
Sequence 6007, Ap  
Sequence 13, Appli  
Sequence 1, Appli  
Sequence 119, App  
Sequence 119, App

ALIGNMENTS

RESULT 1  
US-10-017-273A-4  
; Sequence 4, Application US/10017273A  
; Publication No. US20030119714A1  
; GENERAL INFORMATION:  
; APPLICANT: Pfizer Inc.  
; APPLICANT: Naylor, Alasdair M.  
; APPLICANT: Van Der Graaf, Pieter H  
; APPLICANT: Wayman, Christopher P.  
; TITLE OF INVENTION: Treatment of Male Sexual Dysfunction  
; FILE REFERENCE: PC22013  
; CURRENT APPLICATION NUMBER: US/10/017,273A  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/265,358  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: GB 0030647.2  
; PRIOR FILING DATE: 2000-12-15  
; PRIOR APPLICATION NUMBER: GB 0108730.3  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: GB 0120679.6  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 09/905,846  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/291,722  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: US 09/895,367  
; PRIOR FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 2893  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-017-273A-4

Query Match 83.2%; Score 15.8; DB 9; Length 2893;  
Best Local Similarity 89.5%; Pred. No. 41;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 GATCGGCTACCTGACTAC 19  
||||| |||||||||  
Db 1705 GATCGGCACCTGACTAC 1723

RESULT 2  
US-09-905-846-1  
; Sequence 1, Application US/09905846  
; Patent No. US20020102707A1  
; GENERAL INFORMATION:  
; APPLICANT: Ian Dennis Harrow  
; APPLICANT: Peter Stacey  
; APPLICANT: Rogerick Thomas Walsh  
; TITLE OF INVENTION: Compounds for the treatment of sexual dysfunction  
; FILE REFERENCE: PCS10926APME  
; CURRENT APPLICATION NUMBER: US/09/905,846  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 0017387.2  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/220,908  
; PRIOR FILING DATE: 2000-07-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2893  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-905-846-1

Query Match 83.2%; Score 15.8; DB 10; Length 2893;  
Best Local Similarity 89.5%; Pred. No. 41;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GATCGGCTACCCCTGACTAC 19  
||||| |||||||||  
Db 1705 GATCGGCACCCCTGACTAC 1723

RESULT 3  
US-10-017-273A-5  
; Sequence 5, Application US/10017273A  
; Publication No. US20030119714A1  
; GENERAL INFORMATION:  
; APPLICANT: Pfizer Inc.  
; APPLICANT: Naylor, Alasdair M.  
; APPLICANT: Van Der Graaf, Pieter H  
; APPLICANT: Wayman, Christopher P.  
; TITLE OF INVENTION: Treatment of Male Sexual Dysfunction  
; FILE REFERENCE: PC22013  
; CURRENT APPLICATION NUMBER: US/10/017,273A  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/265,358  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: GB 0030647.2  
; PRIOR FILING DATE: 2000-12-15  
; PRIOR APPLICATION NUMBER: GB 0108730.3  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: GB 0120679.6  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 09/905,846  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/291,722  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: US 09/895,367  
; PRIOR FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 2975  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-017-273A-5

Query Match 83.2%; Score 15.8; DB 9; Length 2975;  
Best Local Similarity 89.5%; Pred. No. 41;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GATCGGCTACCCCTGACTAC 19  
||||| |||||||||  
Db 1770 GATCGGCACCCCTGACTAC 1788

RESULT 4  
US-09-905-846-5  
; Sequence 5, Application US/09905846  
; Patent No. US20020102707A1  
; GENERAL INFORMATION:  
; APPLICANT: Ian Dennis Harrow  
; APPLICANT: Peter Stacey  
; APPLICANT: Rogerick Thomas Walsh  
; TITLE OF INVENTION: Compounds for the treatment of sexual dysfunction  
; FILE REFERENCE: PCS10926APME  
; CURRENT APPLICATION NUMBER: US/09/905,846  
; CURRENT FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 0017387.2  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/220,908  
; PRIOR FILING DATE: 2000-07-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 2975  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-905-846-5

Query Match 83.2%; Score 15.8; DB 10; Length 2975;  
Best Local Similarity 89.5%; Pred. No. 41;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GATCGGCTACCCCTGACTAC 19  
||||| |||||||||  
Db 1770 GATCGGCACCCCTGACTAC 1788

RESULT 5  
US-09-938-842A-2328/c  
; Sequence 2328, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAIN:  
; FILE REFERENCE: SCRIPL300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2328  
; LENGTH: 432  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2328

Query Match 81.1%; Score 15.4; DB 9; Length 432;  
Best Local Similarity 94.1%; Pred. No. 74;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCGGCTACCCCTGACTAC 19  
||||| |||||||||  
Db 140 TCGGCTGCCCTGACTAC 124

```

RESULT 7
US-09-770-445-294/c
; Sequence 294, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 202305 (PARA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770,445
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 294
; LENGTH: 967
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-445-294

```

QY 1 GATCGGCTACCTGACTA 18  
||||| |||||  
db 59 GATCGGCTCCCTGACCA 42





Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATCGGCTACCGTGAC 16  
||||| |||||  
Db 541 GATCGGCGACCGTGAC 526

## RESULT 15

US-10-037-270-156/c  
; Sequence 156, Application US/10037270  
; Publication No. US20030104529A1

## ; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chinghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunqing  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Tillinghast, John  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and  
; FILE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/10/037, 270  
; CURRENT FILING DATE: 2002-01-04  
; PRIOR APPLICATION NUMBER: 09/552, 317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488, 725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1104  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 156  
; LENGTH: 4053  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (409)..(3819)  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(4053)  
; OTHER INFORMATION: n = a,t,c or g

US-10-037-270-156

Query Match 75.8%; Score 14.4; DB 9; Length 4053;  
Best Local Similarity 93.8%; Pred. No. 2.2e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATCGGCTACCGTGAC 16  
||||| |||||  
Db 2742 GATCGGCTCCCGTGAC 2727

Search completed: July 9, 2003, 02:22:37  
Job time : 113.787 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 00:47:28 ; Search time 963.254 Seconds  
(without alignments)  
319.453 Million cell updates/sec

Title: US-09-647-780A-19  
Perfect score: 19  
Sequence: 1 gatcgctaccctgactac 19

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hnc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hnc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pin:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.4	86.3	268	BF798255	BF798255 RC4-C1011
2	16.4	86.3	564	BE517943	BE517943 WHE0804.A
3	16.4	86.3	580	BE444890	BE444890 WHE1129.C
4	16.4	86.3	667	BJ283075	BJ283075 BJ283075
5	15.8	83.2	261	BQ252570	BQ252570 san77f05.
6	15.8	83.2	304	D22128	D22128.RICC10388A

7	15.8	83.2	335	9	AU068505
8	15.8	83.2	360	17	AZ860366
9	15.8	83.2	362	9	AU082608
10	15.8	83.2	377	9	AA965735
11	15.8	83.2	396	9	AL837034
12	15.8	83.2	439	9	AI978418
13	15.8	83.2	505	12	BF009157
14	15.8	83.2	659	10	BB279096
15	15.8	83.2	851	12	BE888846
16	15.8	83.2	914	10	AW155204
17	15.4	81.1	247	14	R84027
18	15.4	81.1	251	14	Z29058
19	15.4	81.1	262	14	BQ459214
20	15.4	81.1	271	10	BA444927
21	15.4	81.1	282	14	R30332
22	15.4	81.1	283	9	AA712729
23	15.4	81.1	292	10	AV562561
24	15.4	81.1	350	14	T21946
25	15.4	81.1	354	14	T22584
26	15.4	81.1	359	10	AV526345
27	15.4	81.1	368	10	BB870158
28	15.4	81.1	373	12	BF046319
29	15.4	81.1	375	14	R90649
30	15.4	81.1	375	14	T44891
31	15.4	81.1	376	14	T41898
32	15.4	81.1	378	14	T21763
33	15.4	81.1	381	9	AA042552
34	15.4	81.1	406	14	T44972
35	15.4	81.1	423	10	AV818992
36	15.4	81.1	424	10	AV531738
37	15.4	81.1	426	14	R64763
38	15.4	81.1	427	10	AV530491
39	15.4	81.1	428	14	T75790
40	15.4	81.1	429	12	BF419565
41	15.4	81.1	432	10	AA463309
42	15.4	81.1	439	10	BE470647
43	15.4	81.1	442	14	H77177
44	15.4	81.1	444	14	T88437
45	15.4	81.1	452	14	R84020

## ALIGNMENTS

RESULT 1  
BF798255  
LOCUS RC4-C10118-021000-011-b02 C10118 Homo sapiens cDNA, mRNA sequence.  
EST 12-JAN-2001  
DEFINITION RC4-C10118-021000-011-b02 C10118 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF798255  
VERSION BF798255.1 GI:12127244  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 268)  
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000).  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC4&st2=RC4-C10118-021000-011-b02&t3=2000-10-02&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 81  
 High quality sequence stop: 268.

#### FEATURES

source  
 1 268  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="C10118"  
 /dev\_stage="Adult"

/note="Organ: colon; ins: Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

66 a 101 c 31 g 70 t

#### BASE COUNT

ORIGIN

Query Match 86.3%; Score 16.4; DB 12; Length 268;  
 Best Local Similarity 94.4%; Pred. No. 8.2e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATCGCTACCTGACTA 18

||||| ||||||| |||||||

Db 52 GATCTGCTACCTGACTA 69

#### RESULT 2

BE517943

LOCUS

DEFINITION  
 WHE0804\_A04\_B08ZS Wheat vernalized crown cDNA library Triticum aestivum cDNA clone WHE0804\_A04\_B08, mRNA sequence.

ACCESSION

BE517943

VERSION

BE517943.1

KEYWORDS

EST.

SOURCE

ORGANISM

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.

1 (bases 1 to 564)

Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Seaton, C.L. and Tong, J.C.

The structure and function of the expressed portion of the wheat genomes - vernalized crown cDNA library

Unpublished (2000)

Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oanderson@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20

Seq primer: Stratagene SK primer.

Location/Qualifiers

1 564

/organism="Triticum aestivum"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone\_lib="WHE0804\_A04\_B08"

/clone\_type="Wheat vernalized crown cDNA library"

/tissue\_type="Crown tissue of seedling"

/dev\_stage="Five-week old seedling"

/lab\_host="E. coli SOLR"

#### FEATURES

source

/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site\_1: EcoRI; Site\_2: XhoI; Seeds were germinated and grown at 4 C for 5 weeks. The tissue, total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give phaluescript phagemids in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 120 a 156 c 161 g 127 t

#### ORIGIN

Query Match 86.3%; Score 16.4; DB 10; Length 564;

Best Local Similarity 94.4%; Pred. No. 1.1e+03;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATCGGCTACCTGACTAC 19

||||| ||||||| |||||||

Db 500 ATCGGCTACCTGACTAC 517

#### RESULT 3

BE444890

LOCUS

DEFINITION  
 WHE1129\_C05\_E09ZS Wheat etiolated seedling root normalized cDNA library Triticum aestivum cDNA clone WHE1129\_C05\_E09, mRNA sequence.

ACCESSION

BE444890

VERSION

BE444890.1

KEYWORDS

EST.

SOURCE

ORGANISM

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.

1 (bases 1 to 580)

Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Nguyen, H.T., Rausch, C.J., Seaton, C.L., Tong, J.C. and Zhang, D.

The structure and function of the expressed portion of the wheat genomes - Normalized root cDNA library

Unpublished (2000)

Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oanderson@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20

Seq primer: Stratagene SK primer.

Location/Qualifiers

1 580

/organism="Triticum aestivum"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone\_lib="WHE1129\_C05\_E09"

/clone\_type="Wheat etiolated seedling root normalized cDNA library"

/tissue\_type="Root"

/dev\_stage="Five day old etiolated seedling"

/lab\_host="E. coli DH108"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid pBluescript SK; Site\_1: EcoRI; Site\_2: XhoI; Seeds were surface-sterilized, germinated and grown aseptically in the dark at room temperature on filter paper with water, nystatin and cefotaxime in covered crystallization dishes. Roots were harvested. The tissue, total RNA, and poly(A) RNA were prepared, a cDNA library was made in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. The cDNA clones were in vivo excised to give pBluescript phagemids before

normalization was carried out. The mass excision of phagemid library and normalization were done in HT Nguyen lab by D. Zhang at Texas Tech University. Normalization protocol used was that of Soares. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 129 a 155 c 162 g 134 t

ORIGIN

Query Match 86.3%; Score 16.4; DB 10; Length 580;

Best Local Similarity 94.4%; Pred. No. 1.1e+03; Mismatches 0; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATCGGCTACCTGACTAC 19

||||| |||||||

DB 472 ATCGGCTACCTGACTAC 489

RESULT 4

BJ283075/c

LOCUS BJ283075 667 bp mRNA linear EST 09-APR-2002

DEFINITION BJ283075 Y. Ogiwara unpublished cDNA library, WH\_r Triticum aestivum cDNA clone whrl4d05 3', mRNA sequence.

ACCESSION BJ283075

VERSION BJ283075.1 GI:20104047

KEYWORDS EST

SOURCE bread wheat.

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae

; Triticeae; Triticum.

1 (bases 1 to 667)

Ogiwara, Y. and Murai, K.

Expressed genes in Triticum aestivum

Unpublished (2002)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

FEATURES

source

1..667

/organism="Triticum aestivum"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="whrl4d05"

/tissue\_type="root"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;

Site 1: EcoRI; Site 2: XhoI; Plants were grown under

hydroponic conditions at UC Davis, salt stressed for 12

hours, and for 7 days, then dissected and frozen (Akhunov

in J. Dvorak Lab). Total RNA was prepared from sheath

tissue, equal quantities of RNA were pooled from the two

samples, polyA was purified from the pooled RNA, a cDNA

library was made, and the cDNA clones were in vivo

excised to give pBluescript phagemids in the TJ Clonase lab

at the University of California, Riverside (Akhunov, Chin

, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).

Plasmid DNA preparations and DNA sequencing were

performed in the OD Anderson lab (all other authors)."

BASE COUNT 147 a 192 c 179 g 149 t

ORIGIN

Query Match

Best Local Similarity 86.3%; Score 16.4; DB 13; Length 667;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATCGGCTACCTGACTAC 19

||||| |||||||

DB 91 ATCGGCTACCTGACTAC 74

RESULT 5

BJ252570

LOCUS

DEFINITION

BJ252570 261 bp mRNA linear EST 03-MAY-2002

San77f05.Y1 Gm-cl052 Glycine max cDNA clone SOYBEAN CLONE ID:

Gm-cl052-5650 5' similar to TR:Q9SX30 Q9SX30 F24J5.9. ; mRNA

sequence.

ACCESSION BJ252570

VERSION BJ252570.1 GI:20448446

KEYWORDS EST

SOURCE soybean.

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 261)

Shoemaker, R., Keim, P., Vodkin, L., Erpelidng, J., Coryell, V., Khanna

, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,

Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers

, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk

, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact: cu@resgen.com web site:

www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 183.

Location/Qualifiers

1..261

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl052-5650"

/clone\_lib="Gm-cl052"

/tissue\_type="whole seedlings of greenhouse grown plants"

/dev\_stage="1 week old"

/lab\_host="DH10B"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; The Harosoy NIL was constructed and seed was

provided by Dr. J. Specht, University of Nebraska

(Shoemaker and Specht, 1995). The cDNA library was

constructed from mRNA isolated from whole seedlings of 1

week old greenhouse grown plants. Complementary DNA was

synthesized from mRNA using a primer consisting of a

poly(dT) sequence with a XhoI restriction site and a 3'

anchor. EcoRI adapters were ligated to the blunt-ended

cDNA fragments followed by XhoI digestion. The cDNA

fragments were directionally cloned into the EcoRI-XhoI

restriction site of the pBluescript vector. The ligated

cDNA fragments were transformed into DH10B host cells

(GibcoBRL). The library was constructed in cooperation

with Dr. Paul Keim's laboratory at Northern Arizona

University."

BASE COUNT 75 a 55 c 58 g 73 t

ORIGIN

Query Match

Best Local Similarity 83.2%; Score 15.8; DB 14; Length 261;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GATCGGCTACCTGACTAC 19

(<http://www.jax.org/resources/documents/dnares/>). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

BASE COUNT 94 a 82 c 71 g 113 t  
 ORIGIN  
 Query Match 83.2%; Score 15.8; DB 17; Length 360;  
 Best Local Similarity 89.5%; Pred. No. 1.8e+03;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GATCGGCTACCCGTGACTAC 19  
 Db 324 GATCGGCTACCCGTGACTAC 342

RESULT 9  
 AU082608 362 bp mRNA linear EST 02-APR-2002  
 LOCUS  
 DEFINITION AU082608 Rice callus Oryza sativa (japonica cultivar-group) cDNA  
 ACCESSION AU082608  
 VERSION AU082608.1 GI:7134660  
 KEYWORDS EST.  
 SOURCE Oryza sativa (japonica cultivar-group).  
 ORGANISM Oryza sativa (japonica cultivar-group).  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Eriophytidae; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzoideae; Oryza.  
 REFERENCE 1 (bases 1 to 362)  
 AUTHORS Sasaki, T. and Yamamoto, K.  
 TITLE Rice cDNA from callus (2000)  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Takuji Sasaki  
 National Institute of Agrobiological Resources  
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
 305-8602, Japan  
 Tel: 81-298-38-7441  
 Fax: 81-298-38-7468  
 Email: tsasaki@abrr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
 PROJECT "RGP"

PROJECT "RGP", Location/Qualifiers

FEATURES  
 source  
 1..362  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="C30226"  
 /clone\_lib="Rice callus"  
 /note="Vector: pBluescript II SK+; Site\_1: SalI; Site\_2: NotI; cDNA prepared from rice callus mRNAs by using oligo(dT) as a primer and ligating to the SalI-NotI site of pBluescript II SK+ phagemid."  
 BASE COUNT 109 a 75 c 98 g 80 t

Query Match 83.2%; Score 15.8; DB 9; Length 362;  
 Best Local Similarity 89.5%; Pred. No. 1.8e+03;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GATCGGCTACCCGTGACTAC 19  
 Db 208 GATCGGCTACCCGTGACTAC 226

RESULT 10  
 AA965735/c  
 LOCUS

DEFINITION AA965735 377 bp mRNA linear EST 31-JUL-1998  
 O4g04al.r1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone  
 O4g04al 5', mRNA sequence.

ACCESSION AA965735.1 GI:3139619  
 VERSION  
 KEYWORDS  
 SOURCE

ORGANISM

Emericella nidulans.

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 Eurotiales; Trichocomaceae; Emericella.

REFERENCE 1 (bases 1 to 377)

AUTHORS Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R.,  
 Prade, R. and Roe, B.

Unpublished (1998)

COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu  
 Department of Chemistry and Biochemistry  
 Advanced Center for Genome Technology, University of Oklahoma  
 620 Parrington Oval, Norman, OK 73019, USA  
 Tel: 405 325 4912  
 Fax: 405 325 7762  
 Email: broe@ou.edu

FEATURES  
 source

1..377  
 Location/Qualifiers

/organism="Emericella nidulans"

/strain="FGSC A26"

/db\_xref="taxon:162425"

/clone="O4g04al"

/clone\_lib="Aspergillus nidulans 24hr asexual  
 developmental and vegetative cDNA lambda zap library"

/tissue\_type="vegetative mycelia, asexual structures"

/note="Vector: pBluescript SK+; Site\_1: EcoRI; Site\_2:  
 XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript  
 3' end of cDNA cloned into XhoI site of pBluescript"

BASE COUNT 79 a 97 c 109 g 92 t

ORIGIN

Query Match 83.2%; Score 15.8; DB 9; Length 377;  
 Best Local Similarity 89.5%; Pred. No. 1.8e+03;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GATCGGCTACCCGTGACTAC 19

Db 134 GATCGGCTACCCGTGACTAC 116

RESULT 11

AL837034

LOCUS

DEFINITION AL837034 EBRF Takifugu rubripes cDNA clone EBRF027apc22, mRNA

ACCESSION AL837034

VERSION AL837034.1 GI:21878996

KEYWORDS

SOURCE

ORGANISM

Takifugu rubripes.

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 Tetraodontidae; Takifugu.  
 1 (bases 1 to 396)  
 Clark, M.S.  
 Takifugu rubripes ESTs  
 Unpublished (2002)

## COMMENT

Contact: Clark MS  
MRC Human Genome Mapping Project Resource Centre  
Hinxton, Cambridge, CB10 1SB, UK  
Email: biohelp@hgmrc.mrc.ac.uk  
Email: biohelp@hgmrc.mrc.ac.uk  
Vector: pBluescript II KS  
V\_type: phagemid  
PRIMER: KS

The clone can be obtained from [www.hgmrc.mrc.ac.uk](http://www.hgmrc.mrc.ac.uk)  
Library created by Greg Elgar  
MRC Human Genome Mapping Project Resource Centre, Hinxton,  
Cambridge, CB10 1SB, UK  
Library sequenced by Sarah Warner and Jim Hills  
MRC Human Genome Mapping Project Resource Centre, Hinxton,  
Cambridge, CB10 1SB, UK.

## FEATURES

Location/Qualifiers

1..396  
/organism="Takifugu rubripes"  
/db\_xref="taxon:31033"  
/clone="EFrf027apcp22"  
/clone\_lib="EFrf"  
/tissue\_type="brain"  
/note="Vector: pBluescript II KS"

BASE COUNT  
ORIGIN

96 a 105 c 94 g 101 t

Query Match 83.2%; Score 15.8; DB 9; Length 396;  
Best Local Similarity 89.5%; Pred. No. 1.9e+03;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GATCGGTACCCCTGACTAC 19  
||||| ||||||| ||  
Db 52 GATCGGTACCCCTGACAAC 70

## RESULT 12

AI978418

LOCUS

DEFINITION AI978418 439 bp mRNA linear EST 19-JAN-2001  
clone R2530, R EcoRI Rice Etiolated Leaf cDNA Library Oryza sativa cDNA

AI978418 1 GI:5803448

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI978418

DEFINITION

AI978418



94 C 126 g 129 t

Query Match	83.2%;	Score 15.8;	DB 12;	Length 505;
Best Local Similarity	89.5%;	Pred. No. 2.1e+03;		
Matches 17;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

RESULT 14	BB279096	BB279096	659 bp	linear	EST, 24-OCT-2001
LOCUS	BB279096	RIKEN full-length enriched,			
DEFINITION		adult retina Mus musculus cDNA			
		clone A930015C18 3', mRNA sequence.			

**SOURCE** house mouse.

REFERENCE 1 (bases 1 to 659)

**TITLE** RIKEN Mouse ESTs (Arakawa, T., et al. 2001).

Contact: Yoshihide Hayashizaki

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Email: [genome-res@qsc.riken.go.jp](mailto:genome-res@qsc.riken.go.jp),

M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y:

genes. Genome Res. 10 (10), 1617-1630 (2000)

S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

Y. and Hayashizaki, Y.

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001).

1. The first step in the process of creating a new product is to identify a market need. This involves conducting market research to understand what consumers want and what problems they are facing. Once a need is identified, the next step is to develop a concept that addresses this need. This is often done through brainstorming sessions with a team of designers and engineers. The concept is then refined through prototyping and testing, ensuring that it meets the requirements of the market. Finally, the product is launched and its performance is monitored to ensure it continues to meet the needs of the market.

cdNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

1

10

10

1

Age Group	Percentage
18-29	85%
30-49	80%
50-69	75%
70+	70%

BASE COUNT

Quarry Mat +

### Best Local Matches

Qy .

Db . 38

DECEMBER 1956

LOCUS  
BE88846/C

2000

VERSION

SOURCE  
ORGANIZATION

## AUTHORS

**JOURNAL**

10

10

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM9737 row: d column: 17

High quality sequence stop: 683.

#### FEATURES

source

location/Qualifiers

1..851

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:3914848"

/clone\_lib="NIH\_MGC\_71"

/tissue\_type="leiomyosarcoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site.1: NotI;

Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 2.1 kb."

184 a 234 c 281 g 152 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 83.2%; Score 15.8; DB 12; Length 851;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GATCGGCTACCTGACTAC 19

||||| ||||| |||||

Db 724 GATCGGCCACCTGCCTAC 706

Search completed: July 8, 2003, 09:22:13

Job time : 967.404 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: July 8, 2003, 00:45:53 ; Search time 209.093 Seconds  
(without alignments)  
2644.537 Million cell updates/sec

Title: US-09-647-780A-20

Perfect score: 19

Sequence: 1 gttcgccatccagtcctac 19

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	19	100.0	19	6	AX014720 Sequence.
2	19	100.0	2765	6	AX014701 Sequence
3	17.4	91.6	2076	6	AX146976 Sequence
4	17.4	91.6	2232	6	AX319864 Sequence
5	17.4	91.6	2262	6	AX146980 Sequence
6	17.4	91.6	2340	6	AX146978 Sequence
7	17.4	91.6	2340	6	AX473102 Sequence
8	17.4	91.6	2583	10	AF302075 Mus muscu
9	17.4	91.6	2601	10	AF157106 Mus muscu
10	17.4	91.6	2636	6	AX139743 Sequence
11	17.4	91.6	2652	10	AF302076 Mus muscu
12	17.4	91.6	2663	6	AX139745 Sequence
13	17.4	91.6	2676	6	AX033274 Sequence
14	17.4	91.6	2694	10	AF302077 Mus muscu
15	17.4	91.6	2714	6	AX139741 Sequence
16	17.4	91.6	2784	9	AF336981 Homo sapi
17	17.4	91.6	2850	9	AK093058 Homo sapi
18	17.4	91.6	2892	10	AF157105 Mus muscu
19	17.4	91.6	2893	6	AX356951 Sequence
20	17.4	91.6	2893	6	AX463057 Sequence
21	17.4	91.6	2925	6	AX033272 Sequence
22	17.4	91.6	2925	10	AF176569 Mus muscu
23	17.4	91.6	2953	6	AX473100 Sequence
24	17.4	91.6	2975	6	AX356955 Sequence
25	17.4	91.6	2975	6	AX463058 Sequence
26	17.4	91.6	121499	2	AC130119 Rattus no
27	17.4	91.6	144820	2	AC099556 Trypanoso
28	17.4	91.6	154736	9	AL139246 Human DNA
29	17.4	91.6	208249	2	AL607032 Mus muscu
30	17.4	91.6	235760	2	AC114403 Mus muscu
31	16.4	86.3	436	10	AF012882 Mus muscu
32	16.4	86.3	456	10	AF012881 Mus muscu
33	16.4	86.3	1096	10	BC023402 Mus muscu
34	16.4	86.3	1431	6	AX418848 Sequence
35	16.4	86.3	1448	10	AF294790 Mus muscu
36	16.4	86.3	1483	10	SC09989 NKX-5.2-NK-
37	16.4	86.3	8401	9	AC009980 Homo sapi
38	16.4	86.3	17941	8	CRE00411X Chlamydomon
39	16.4	86.3	36758	2	AC017690 Drosophil
40	16.4	86.3	87817	9	AC079174 Homo sapi
41	16.4	86.3	88647	2	AC114476 Magnaport
42	16.4	86.3	98151	2	AC119404 Magnaport
43	16.4	86.3	129274	2	AC095509 Rattus no
44	16.4	86.3	132906	2	AC109802 Canis fam
45	16.4	86.3	144426	2	AP004854 Oryza sat

# ALIGNMENTS

RESULT 1	AX014720	AX014720	Sequence 20 from Patent WO9953077.	19 bp	DNA	linear	PAT 07-SEP-2000
LOCUS	AX014720	Sequence 20 from Patent WO9953077.					
DEFINITION	AX014720						
ACCESSION	AX014720						
VERSION	AX014720.1	GI:10040993					
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							

Pred. No. is the number of results predicted by chance to have a

```

Db      2281 GTTCGCATCCAGTCCATC 2299
|||||
RESULT 3
AX146976
LOCUS   AX146976
DEFINITION Sequence 1 from Patent WO0136610.
ACCESSION AX146976
VERSION  1
KEYWORDS AX146976.1 GI:14346247
SOURCE  human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2076)
AUTHORS  Deleersnijder,W., Wiegers,R. and Weske,M.
TITLE    Human enzymes of the metalloprotease family
JOURNAL  Patent: WO 0136610-A 1 25-MAY-2001;
          Solvay Pharmaceuticals B.V. (NL)
FEATURES
          source
            1..2076
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              <i>1..2076
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAC41159.1"
                /db_xref="GI:14346248"
                /translation="CTTGCVIAAIRLQMDTTEPCDDFYOFACGGWLRHVVIPET
                NGRSYFDVFLRDEVLVILKAVLENSAKORPAVEKARTLYRSCMNQSVLEKRGSQPLL
                DILEVYGGWVPAWRNWTGVLGEMELERQALMNSQFNRRVLIDLFIWDDQNSRRHI
                IYIDQTLGMPREYTFNGSNNKREAVLQFMVSVATILLREDANLPRDSCLVQEDMM
                OVLEDTOLAKATVPOEERHDVITLYHRMGLLELOSQFGLKGFENWTLFIOTVLSSVKI
                KLLPDEEVVYGPYIQLNLENIIDTYSARTIONLYVRRLVLRIGSLRSORFKUTRVNY
                RKALPDEVEEVWRCVQVYVSNMENVAGSLVRYEAPFGDSKSMVRELIDKRVTFVFE
                DLELGMWDESKKKAQKRAMSIREIQIHPDYLIEEMNRRLDEEYSLNLFSDLFIE
                NSLQNLKVGAKRSRLKRLEKVDNLWIIIGAAVYNFYSPNRNQIVTPAGILQFPFFSK
                EQPALNGFGIGMIVGHEITHGFDNGRNFDRKNGNMMDMWSNFSTQHFREQSECMYFO
                YGNYSWDLADEONVNGFNTLGENIADNGGVQRAYKAYLKWMAEGGKQQQLPGLDLTHE
                QLFIIYNAOVCGWSYRPEAFQISIKTDVHSPLKRVILGSLQNLAAAPADPTFCARGTDM
                HPKRCRVW"
          BASE COUNT  493 a      578 c      636 g      369 t
          ORIGIN
            Query Match          91.6%; Score 17.4; DB 6; Length 2076;
            Best Local Similarity 94.7%; Pred. No. 5e+02;
            Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GTTCGCATCCAGTCCATC 19
          |||||||
Db      1926 GTTCGCATCCAGTCCATC 1944

RESULT 4
AX319864
LOCUS   AX319864
DEFINITION Sequence 28 from Patent WO0183782.
ACCESSION AX319864
VERSION  1
KEYWORDS AX319864.1 GI:17901454
SOURCE  human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS  Ploeman,G.D., Whyte,D., Sudarsanam,S., Manning,G., Caenepeel,S. and
          Payne,V.
TITLE    Novel proteases
JOURNAL  Patent: WO 0183782-A 28 08-NOV-2001;
          Sugen, Inc. (US)
FEATURES
          source
            1..2232
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              <i>1..2232
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAC41159.1"
                /db_xref="GI:14346248"
                /translation="CTTGCVIAAIRLQMDTTEPCDDFYOFACGGWLRHVVIPET
                NGRSYFDVFLRDEVLVILKAVLENSAKORPAVEKARTLYRSCMNQSVLEKRGSQPLL
                DILEVYGGWVPAWRNWTGVLGEMELERQALMNSQFNRRVLIDLFIWDDQNSRRHI
                IYIDQTLGMPREYTFNGSNNKREAVLQFMVSVATILLREDANLPRDSCLVQEDMM
                OVLEDTOLAKATVPOEERHDVITLYHRMGLLELOSQFGLKGFENWTLFIOTVLSSVKI
                KLLPDEEVVYGPYIQLNLENIIDTYSARTIONLYVRRLVLRIGSLRSORFKUTRVNY
                RKALPDEVEEVWRCVQVYVSNMENVAGSLVRYEAPFGDSKSMVRELIDKRVTFVFE
                DLELGMWDESKKKAQKRAMSIREIQIHPDYLIEEMNRRLDEEYSLNLFSDLFIE
                NSLQNLKVGAKRSRLKRLEKVDNLWIIIGAAVYNFYSPNRNQIVTPAGILQFPFFSK
                EQPALNGFGIGMIVGHEITHGFDNGRNFDRKNGNMMDMWSNFSTQHFREQSECMYFO
                YGNYSWDLADEONVNGFNTLGENIADNGGVQRAYKAYLKWMAEGGKQQQLPGLDLTHE
                QLFIIYNAOVCGWSYRPEAFQISIKTDVHSPLKRVILGSLQNLAAAPADPTFCARGTDM
                HPKRCRVW"
          BASE COUNT  493 a      578 c      636 g      369 t
          ORIGIN
            Query Match          91.6%; Score 17.4; DB 6; Length 2076;
            Best Local Similarity 94.7%; Pred. No. 5e+02;
            Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

PAT 08-JUN-2001
PAT 14-DEC-2001

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
512 a 620 c 705 g 395 t

Query Match
Best Local Similarity 91.6%; Score 17.4; DB 6; Length 2232;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19
|||||
Db 2082 GTTCGCCATCCATCCATC 2100

RESULT 5
AX146980
LOCUS AX146980
DEFINITION Sequence 5 from Patent WO0136610.
ACCESSION AX146980
VERSION AX146980.1 GI:14346251
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2262)
AUTHORS Deleersnijder W., Wiegers, R. and Weske, M.
TITLE Human enzymes of the metalloprotease family.
JOURNAL Patent: WO 0136610-A 5 25-MAY-2001;
Solvay Pharmaceuticals B.V. (NL)
FEATURES
source
1..2262
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..2262
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAC41161.1"
/db_xref="GI:14346251"
/translation="MGKSEGPVGMVESAGRAGOKRPGFLEGLLLLLLLVTAALVALG
VLYADRRGIPAEQSEVCTTPGCVIAAARILQNMPTTEPCDDYFQACGWLRRHV
IPENSRYSIFDVLRDELEILKAVLENSTAKDRPAVEKARTLYRSCMNSVIERKGS
QRLDILLEVGGVPMVMDRWNETVGLWELEQLALMNSQFNRLDLFTWDDQNS
EDMMVLELETQAKATVPOERHDVIALYRMGLEELQSQFGLKGFNTLFIQTVL
SVKIKLPLDEEVVYGYPIYQLNENIIDTYSARTIONYLWRLVLDIGLSQRFKDT
RVNYRKALFTGMVEVRWECVGVNSMENAVGSLYVREAFPGDSKSMYRELIDKVR
TVFVETLDELGMDESKKRAQEKAMSIREQIGHDPYILEEMNRRLDEEYSLNASED
LYFENSQNLKVGAKRSRLKREKVDPLNLTIGAAVNAFYSPNENIIVFPAGLQPP
FFSKQEQALNFGGIGMVGHEITHGFDNDRNFDKNGMMDWNSNSTOHPRESEC
MTYQYGNYSWDLADEQNVNGFNTLGNADGVGRQAKYKAYLKWNAEGGKDDQLPGLD
LTHEQLFFINAYOVWCGSYRPEFAIQSIKTDVHSPLKYRVLGSLQNLAAAFADTFHCR
GTPMHPKRCRW"
BASE COUNT 520 a 628 c 716 g 398 t
ORIGIN

Query Match
Best Local Similarity 91.6%; Score 17.4; DB 6; Length 2262;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19
|||||
Db 2112 GTTCGCCATCCATCCATC 2130

RESULT 6
AX146978
LOCUS AX146978
DEFINITION Sequence 3 from Patent WO0136610.
ACCESSION AX146978
VERSION AX146978.1 GI:14346249
KEYWORDS
SOURCE human.

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2340)
AUTHORS Deleersnijder W., Wiegers, R. and Weske, M.
TITLE Human enzymes of the metalloprotease family
JOURNAL Patent: WO 0136610-A 3 25-MAY-2001;
Solvay Pharmaceuticals B.V. (NL)
FEATURES
source
1..2340
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
<1..2340
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAC41160.1"
/db_xref="GI:14346250"
/translation="MGKSEGPVGMVESAGRAGOKRPGFLEGLLLLLLLVTAALVALG
VLYADRRGKQLPRLASRLCFLEERTFKRPRGIPAEQSEVCTTPGCVIAAARIL
QNMPTTEPCDDYFQACGWLRRHVIPENSRYSIFDVLRDELEILKAVLENSTAK
DRPAVEKARTLYRSCMNSVIERKGSQPLDILEVGGVPMVMDRWNETVGLWELE
QLALMNSQFNRLDLFTWDDQNSRRHIYIDOPTLCMPSRVYENGSRKVRREA
YLOFMVSVATLLREDANLPRDSCVQSDMMQVLETLAKATVPOERHDVIALYHR
MGLEELQSQFGLKGFNTLFIQTVLSSVKIKLPLDEEVVYGYPIYQLNENIIDTYS
RTIONYLWRLVLDIGLSQRFKDTRVNYRKALFTGMVEVRWECVGVNSMENNA
VGSILYVREAFPGDSKSMYRELIDKVRTVFVETLDELGMDESKKRAQEKAMSIREQI
GHDPYILEEMNRRLDEEYSLNASEDLYFENSQNLKVGAKRSRLKREKVDPLNLTIG
GAAVNAFYSPNENIIVFPAGLQPPFFSKQEQALNFGGIGMVGHEITHGFDNDR
NFDKNGMMDWNSNSTOHPRESECMIYQYGNYSWDLADEQNVNGFNTLGNADNG
GVROAYKAYLKWNAEGGKDDQLPGLDTHQLFFINAYOVWCGSYRPEFAIQSIKTDV
HSPLKYRVLGSLQNLAAAFADTFHCRGTPMHPKRCRW"
BASE COUNT 539 a 649 c 739 g 413 t
ORIGIN

Query Match
Best Local Similarity 91.6%; Score 17.4; DB 6; Length 2340;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19
|||||
Db 2190 GTTCGCCATCCATCCATC 2208

RESULT 7
AX473102
LOCUS AX473102
DEFINITION Sequence 3 from Patent WO0226958.
ACCESSION AX473102
VERSION AX473102.1 GI:22207836
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Bandaru, R. and Silos-Santiago, I.
TITLE Human neprilysin protease
JOURNAL Patent: WO 0226958-A 3 04-APR-2002;
MILLENNIUM PHARMACEUTICALS, INC. (US)
FEATURES
source
1..2340
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
538 a 647 c 740 g 415 t
BASE COUNT
ORIGIN

Query Match
Best Local Similarity 91.6%; Score 17.4; DB 6; Length 2340;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19
|||||
Db 2190 GTTCGCCATCCATCCATC 2208

```

```

RESULT 8
AF302075      2583 bp      mRNA      linear      ROD 11-JUN-2001
LOCUS        Mus musculus neprilysin-like peptidase alpha mRNA, complete cds.
DEFINITION
ACCESSION    AF302075
VERSION      AF302075.1 GI:10505359
KEYWORDS
SOURCE
ORGANISM     Mus musculus
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS      Shirota, K., Tsubuki, S., Iwata, N., Takaki, Y., Harigaya, W.,
              Maruyama, K., Kiryu-Seo, S., Kiyama, H., Iwata, H., Tomita, T.,
              Iwatsubo, T., and Saido, T.C.
TITLE        Neprilysin degrades both amyloid beta peptides 1-40 and 1-42 most
              rapidly and efficiently among thiorphan- and
              phosphoramidon-sensitive endopeptidases
JOURNAL      J. Biol. Chem. 276 (24), 21895-21901 (2001)
MEDLINE      21293028
PUBMED       11278416
REFERENCE
AUTHORS      Shirota, K. and Saido, T.C.
TITLE        Direct Submission
JOURNAL      Submitted (02-SEP-2000) Proteolytic Neuroscience Laboratory, Brain
              Science Institute, RIKEN, 2-1 Hirotsawa, Wako-shi, Saitama
              351-0198, Japan
FEATURES
Source       Location/Qualifiers
1..2583
   /organism="Mus musculus"
   /db_xref="taxon:10090"
   /note="endopeptidase"
   /product="neprilysin-like peptidase alpha"
   /db_xref="GI:10505360"
   /translation="MVERAGWCKKSPGVEYGLMVLILLGAILVTLGVFYSIALRD
   SSKSDICTPSCVTAARILENMDQSRNPFCEYQACGWLHRHVIPETNSRYSVF
   DIIRDEVLKGLVEDSTSOHRPAVEKAKTLRSCMQSVIEKRDSEPLLSVLRKVG
   GWPVADKWNETMGLKWELEQLAVNSQFNRRVLIDLFINWDDQNSRHVVIDOPT
   LGMPSEYFQEDNNHVKRKALEYFTSVATMLRKDQNLKSAMVREEMAEVLELET
   HLANATVPOEKHRDVTALYHRMDLMELQERFGLKGFNTLFIQNVLSSEVEFLPDEE
   VVYGIPTLENLEDDISYARTMQLVRLVLDKRGISLQSFKEARVDYKALYGT
   TVEEVWRRCVSVNMSAVGSLYIKRAFSDKSTVRELIEKIRSVFVMDLELN
   WDEESKKAQKAMNIREQIGYDYLIDNNKHLDEEYSSLYFFEDLYFENGLQSLN
   NNAORSLKLRKVDQNLWLTIGAAVNAFVSPNRQIVFPAGILOPPFFSKDQPSLN
   FGGIGNVGHEITHGFDNGRNFKNMGLDWSNFSARHFOQSCMTIYQGNFSWE
   LADNQNVGFSTLGENIADNGVRQAKYALRWLADGGDKDQRLPGLNLTVAQLFFIN
   AQWCGSYRPEFAVQSIKTVDHSPKLYRVLGSLQNLPGFSEAFHCPGRSPMPMKRCR
   IW"
BASE COUNT   665 a 667 c 736 g 515 t
ORIGIN
Query Match 91.6%; Score 17.4; DB 10; Length 2583;
Best Local Similarity 94.7%; Pred. No. 5.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCGCCATCCAGTCCATC 19
||||| |||||||
Db 2103 GTTCGCCATCCAGTCCATC 2121

RESULT 9
AF157106      2601 bp      mRNA      linear      ROD 25-NOV-1999
LOCUS        Mus musculus soluble secreted endopeptidase delta mRNA,
DEFINITION    alternatively spliced product, complete cds.
ACCESSION    AF157106
VERSION      AF157106.1 GI:6467400
KEYWORDS
SOURCE
ORGANISM     Mus musculus
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Jagerschmidt, A., Agnel, M. and Culouscou, J.M.
TITLE        Three neprilysin-like membrane metalloproteinases
JOURNAL      Patent: EP 1069188-A 3 17-JAN-2001;
              SANOFI-SYNTHELABO (FR)
FEATURES
Source       Location/Qualifiers
1..2636
   /organism="Mus musculus"
   /db_xref="taxon:10090"
   /note="SEP(delta); metalloprotease; alternatively spliced"
   /codon_start=1
   /product="soluble secreted endopeptidase delta"
   /protein_id="AAF13153.1"
   /db_xref="GI:6467401"
   /translation="MVERAGWCKKSPGVEYGLMVLILLGAILVTLGVFYSIALRD
   SSKSDICTPSCVTAARILENMDQSRNPFCEYQACGWLHRHVIPETNSRYSVF
   DIIRDEVLKGLVEDSTSOHRPAVEKAKTLRSCMQSVIEKRDSEPLLSVLRKVG
   GWPVADKWNETMGLKWELEQLAVNSQFNRRVLIDLFINWDDQNSRHVVIDOPT
   LGMPSEYFQEDNNHVKRKALEYFTSVATMLRKDQNLKSAMVREEMAEVLELET
   HLANATVPOEKHRDVTALYHRMDLMELQERFGLKGFNTLFIQNVLSSEVEFLPDEE
   VVYGIPTLENLEDDISYARTMQLVRLVLDKRGISLQSFKEARVDYKALYGT
   TVEEVWRRCVSVNMSAVGSLYIKRAFSDKSTVRELIEKIRSVFVMDLELN
   WDEESKKAQKAMNIREQIGYDYLIDNNKHLDEEYSSLYFFEDLYFENGLQSLN
   NNAORSLKLRKVDQNLWLTIGAAVNAFVSPNRQIVFPAGILOPPFFSKDQPSLN
   FGGIGNVGHEITHGFDNGRNFKNMGLDWSNFSARHFOQSCMTIYQGNFSWE
   LADNQNVGFSTLGENIADNGVRQAKYALRWLADGGDKDQRLPGLNLTVAQLFFIN
   AQWCGSYRPEFAVQSIKTVDHSPKLYRVLGSLQNLPGFSEAFHCPGRSPMPMKRCR
   IW"
BASE COUNT   655 a 681 c 748 g 517 t
ORIGIN
Query Match 91.6%; Score 17.4; DB 10; Length 2601;
Best Local Similarity 94.7%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTTCGCCATCCAGTCCATC 19
||||| |||||||
Db 2137 GTTCGCCATCCAGTCCATC 2155

RESULT 10
AX139743      2636 bp      DNA      linear      PAT 30-MAY-2001
LOCUS        Sequence 3 from Patent EP1069188.
DEFINITION    AX139743
ACCESSION    AX139743
VERSION      AX139743.1 GI:14275325
KEYWORDS
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Jagerschmidt, A., Agnel, M. and Culouscou, J.M.
TITLE        Three neprilysin-like membrane metalloproteinases
JOURNAL      Patent: EP 1069188-A 3 17-JAN-2001;
              SANOFI-SYNTHELABO (FR)
FEATURES
Source       Location/Qualifiers
1..2636

```

5'UTR  
CDS

/organism="Homo sapiens"

/db\_xref="taxon:9606"

1..7

8..2269

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC39975.1"

/db\_xref="GI:14275326"

/translation="MGKSEGPVGMVESAGRAGOKRPGFLEGGLLLLLLVTAALVALG  
VLYADRRGIPQAQSEVCTTTPGCVIAARILQNDPTTEPCDDPYQACGGVLRVH  
IPENRSYSPFDVLRDELEVLKVGLEDSTQHRPAVEKAKTLYS  
OPLDLILEVGVQVADRWNETVLEWELERQALMNSQFNRRVLIIDLFWDNDONS  
SRHIIYDQPTLGMPSREIYFNGSNRKRREYALQFVSVATLLREDANLPRDSCVQ  
EDMVQVLETLQAKATVQEEHDVIALYHRMGLLELOSOGLKGFDPWTLFIQTVLS  
SVKILLPDEVVYVYGIPIQLNLENIIDTYSARTIONVLRVLDLIGLSQRKDT  
RVNYRKALFTGMVEVRWRECVSVNMSNAVGLSVREAFPGDSKSMVRELIDKVR  
TVFVETDELGMWDEESKKAOKAEKAMNIREQIYGYDYILEDNKKHL  
LYFENSLQNLKVAQSRSLRKLREKVDQNLMTIIGAAVNVAFYSPNRNQTVPFAGILQPP  
FFSKEQPALNFGGIGMVGIEHTHGDGNGRNFKNMMDWNSFSTOHRFQSEC  
MIYQYQNSWDLADEQNVNGFNTIGENIADNGGVROAYKAYLKWMAEGGQQQLPGLD  
LTHEQLFFINVAOVWCGSYRPEFAIQSIKTDVHSPKRYRVLGSLQNLAAFDTHCAR  
GTPMHPKRCRVN"

2270..2635

BASE COUNT 601 a 759 c 813 g 462 t 1 others

ORIGIN

Query Match 91.6%; Score 17.4; DB 6; Length 2636;  
Best Local Similarity 94.7%; Pred. No. 5.2e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GTTCGCCATCCAGTCCATC 19

|||||

Db 2119 GTTCGCCATCCAGTCCATC 2137

RESULT 11

AF302076

LOCUS AF302076

DEFINITION Mus musculus neprilysin-like peptidase beta mRNA, complete cds.

ACCESSION AF302076

VERSION AF302076.1 GI:10505361

KEYWORDS

SOURCE Mus musculus.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

1 (bases 1 to 2652)

Shirotani, K., Kiryu-Seo, S., Kiyama, H., Iwata, H., Tomita, T.,

Maruyama, K., Tsubuki, S., Iwata, N., Takaki, Y., Harigaya, W.,

Iwatsubo, T. and Saido, T.C.

Neprilysin degrades both amyloid beta peptides 1-40 and 1-42 most

rapidly and efficiently among thiorphan- and

phosphoramidon-sensitive endopeptidases

J. Biol. Chem. 276 (24), 21895-21901 (2001)

2 (bases 1 to 2652)

Shirotani, K. and Saido, T.C.

Direct Submission

Submitted (02-SEP-2000) Proteolytic Neuroscience Laboratory, Brain

Science Institute, RIKEN, 2-1 Hirotsawa, Wako-shi, Saitama

351-0198, Japan

FEATURES

Location/Qualifiers

1..2652

/organism="Mus musculus"

/db\_xref="taxon:10090"

25..2322

/note="endopeptidase"

/codon\_start=1

/product="neprilysin-like peptidase beta"

/protein\_id="AAG18447.1"

/db\_xref="GI:10505362"

/translation="MVERAGNCRKSPGFVEGLVLLLLLLGAIIVTLGVFYSIGQL"

PLLTSLHFSWERTVVKRALRDSLSKSDICTTPSCVIAAARILENDDQSRNPNCFY  
OYACGGVLRHVIIPETNSRYSPFDVLRDEVLKVGLEDSTQHRPAVEKAKTLYS  
CMNQSVLEKRSDEPLLSVKMGVGPVADRWNETVLEWELERQALMNSQFNRRVLI  
IDLFINWDDONSRRHVIYIDQPTLGMPSREIYFQEDNNHVKRYKAYLEFVSATMLRK  
DONLSKESAMVREMAVLELETHLANATVPOEKRHDVTALYHRMDLWELQERGLRG  
FNWTLFIQNVLSVELEFPDEEVVYGIPILENLEIDIISYARTIONVLRVLD  
RTGSLSORFEKEARDVYKALYGTTVBEVRWRECVSVNMSNAVGLSVYIKRAFSKDS  
KTSRELIEKIRSVFVNDLNMWDEESKKAOKAEKAMNIREQIYGYDYILEDNKKHL  
DEYSLFTFYDLYFENGLOLNKNAORSLKRLREKVDQNLMTIIGAAVNVAFYSPNRN  
QIVFAGILQPPFFSKDQPSLNFSGIGMVGIEHTHGDGNGRNFKNMMDWNSFSTOHR  
FSARFQOOSQCMYIYQYFNFSWELADNQNVNGFNTIGENIADNGGVROAYKAYLKWMA  
DGKQDRLPLGLNLTAQLPFIINVAOVWCGSYRPEFAIQSIKTDVHSPKRYRVLGSLQNL  
LPGFEAFHCPRGSPMHPKRCRVN"

BASE COUNT 682 a 685 c 755 g 468 t

ORIGIN

Query Match

Best Local Similarity 91.6%; Score 17.4; DB 10; Length 2652;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GTTCGCCATCCAGTCCATC 19

|||||

Db 2172 GTTCGCCATCCAGTCCATC 2190

RESULT 12

AX139745

LOCUS AX139745

DEFINITION Sequence 5 from Patent EP1069188.

ACCESSION AX139745

VERSION AX139745.1 GI:14275327

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

1 (bases 1 to 2653)

Jagerschmidt, A., Agnel, M. and Culoussou, J.M.

Three neprilysin-like membrane metalloproteinases

Patent: EP 1069188-A 5 17-JAN-2001;

SANOFI-SYNTHELABO (FR)

FEATURES

Location/Qualifiers

1..2663

/organism="Homo sapiens"

/db\_xref="taxon:9606"

1..7

8..2296

/note="unnamed protein product"

/codon\_start=1

/db\_xref="GI:14275328"

/translation="MGKSEGPVGMVESAGRAGOKRPGFLEGGLLLLLLVTAALVALG

VLYADRRGKQLPRLASRLCFLOEERTFVKRPRGIPAEQSEVCTTTCVTAARIL

QNDPTTEPCDDFYQFACGGVLRHVIIPETNSRYSPFDVLRDEVLKVALENSTAK

DRPAVEKARTILRSCMNSVLEKRGSOPLDLILEVGGVPMADRWNETVYVQOARPHR

PLHLERPELOPAHLHOPTLGMPSREIYFNGSNRKRREYALQFVSVATLLREDAN

LPDSCVLOEDMVQVLETLQAKATVQEEHDVIALYHRMGLLELOSOGLKGFDPW

TLFTQTVLSSVKIKLLPDEVVYGIPIQLNLENIIDTYSARTIONVLRVLDLDRIG

SLSORFEDTRVNYKALYGTTVBEVRWRECVSVNMSNAVGLSVYIKRAFSKDS

VRELIDKRVFVETDELGMWDEESKKAOKAEKAMNIREQIYGYDYILEDNKKHL

YSLNFSDELIFENSLQNLKVAQSRSLRKLREKVDQNLMTIIGAAVNVAFYSPNRN

QIVFAGILQPPFFSKDQPSLNFSGIGMVGIEHTHGDGNGRNFKNMMDWNSFST

QHRFQSECMIYQYFNFSWELADNQNVNGFNTIGENIADNGGVROAYKAYLKWMAEGG

KDQQLPDLTLTHEQLFFINVAOVWCGSYRPEFAIQSIKTDVHSPKRYRVLGSLQNL

AAFDTHCARGTPMHPKRCRVN"

2294..2663

BASE COUNT 609 a 769 c 816 g 468 t

ORIGIN

1 others

Query Match

Best Local Similarity 91.6%; Score 17.4; DB 6; Length 2663;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;





/translation="MGKSEGPVGVESAGRGOKRPGFLEGGILLLLLVTRALVALG  
VLYADRRGKQLPRLASRLCLOERTFVRKPRGIFEAQEVSEVCTPCGVIAARIL  
QNMDDTTEPCDDFYOFACGGWLRHVVIPETNSRYSDVIRDELEVILKAVLENSTAK  
DRPAVEKARTLYRSCMNOSVIEKRSQPLDILEVVGWMPVAMDRWNETVGLWELER  
OLALANSQFNRRVLIDLFIWDDQSSRHIIYDQPTLGMPSREYFNGSGNRKVKREA  
YLOFVSVATLLREDANLPDRSCLVOEDMVOLETQLAKATVQOERHDVIALYHR  
MGLELQSOGLKGFDTLFIQTVLSSVKILLPDEEVVYGPVYLQNLNIDTYS  
RTIONYLVRVLDRIGLSQRKDTRVNYRKALFGTVEEVWRRECVCYVNSMENA  
VGSLLVREAFPGDSKSMVRELIDKVRTVETLDELGMDEESKKKAQEKAMSIREQI  
GHPDYILEETNRRLDEEYSLNLFSEDLFEENSLQNLKVGAKSLRKLREKVDPNLWII  
GAAVYNAYSPNRNQIVFPAGILQPPFESKEOPQALNFGGIGMVGHEITHGFDNDGR  
NFDXNGNMMDWNSFSTOHFREQSECMYQYGNYSWDLADEQNVNGFNTLGENIADNG  
GVQRAYKAYLKWMAEGGKQQLPGLDLTHEQLFFINYAQVWCGSVRPERFAIQSIKTDV  
HSPILRYVLGSLQNLAAAFADTFHCARGTTPMHPKRCRVW"  
2348..2713

3'UTR  
BASE COUNT 621 a 780 c 836 g 476 t 1 others  
ORIGIN

Query Match 91.6%; Score 17.4; DB 6; Length 2714;  
Best Local Similarity 94.7%; Pred. No. 5.2e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GTTCGCCATCCAGTCCATC 19  
Db ||||| ||||| ||||| |||||  
2197 GTTCGCCATCCATCCATC 2215

Search completed: July 8, 2003, 03:35:10  
Job time : 211.093 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 00:43:28 ; Search time 119.376 Seconds  
(without alignments)  
358.431 Million cell updates/sec

Title: US-09-647-780A-20

Perfect score: 19

Sequence: 1 gttcgccatccagtcacac 19

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_101002.\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	100.0	19	AAZ28826	Rat membrane metal
2	19	100.0	2286	AAZ28130	Soluble secreted e
3	19	100.0	2765	AAZ28810	Rat membrane metal
4	17.4	91.6	324	ABV14605	Human prostate exp
5	17.4	91.6	340	ABV05436	Human prostate exp
6	17.4	91.6	569	ABV35683	Human prostate exp
7	17.4	91.6	569	ABV44488	Human prostate exp
8	17.4	91.6	1124	AAH26248	Human prostate pro
9	17.4	91.6	1226	AAH35007	Human colon cancer

c	10	17.4	91.6	1746	23	ABV23433
c	11	17.4	91.6	1746	23	ABV29290
	12	17.4	91.6	2076	22	AAF89737
	13	17.4	91.6	2232	24	AAF97186
	14	17.4	91.6	2262	22	AAF89739
	15	17.4	91.6	2318	24	AAD30580
	16	17.4	91.6	2340	22	AAF89738
	17	17.4	91.6	2580	24	ABN84280
	18	17.4	91.6	2636	22	AAF59660
	19	17.4	91.6	2663	22	AAF59661
	20	17.4	91.6	2676	21	AAF63764
	21	17.4	91.6	2714	22	AAF59659
	22	17.4	91.6	2893	24	ABN84279
	23	17.4	91.6	2893	24	AAD28544
	24	17.4	91.6	2925	21	AAA63763
	25	17.4	91.6	2953	24	ABK48251
	26	17.4	91.6	2975	24	AAD28547
	27	16.4	86.3	1431	24	ABA99061
c	28	15.8	83.2	398	22	AAF64428
c	29	15.8	83.2	449	21	AAA68155
	30	15.8	83.2	548	23	AAS57710
c	31	15.8	83.2	642	24	ABK75097
c	32	15.8	83.2	649	24	ABN63181
c	33	15.8	83.2	676	22	AAS41344
c	34	15.8	83.2	1328	24	ABK11711
c	35	15.8	83.2	1479	24	ABA05334
	36	15.8	83.2	2446	19	AAV69334
	37	15.8	83.2	2446	21	AA54597
c	38	15.8	83.2	2491	22	AAH17539
c	39	15.8	83.2	2557	24	ABA05333
	40	15.8	83.2	2873	24	ABK70021
	41	15.8	83.2	3474	22	AAH02894
	42	15.8	83.2	3474	22	AAF22449
	43	15.8	83.2	3474	22	ABK70020
c	44	15.8	83.2	10115	22	AAAL05253
c	45	15.8	83.2	10115	23	ABL98136

#### ALIGNMENTS

#### RESULT 1

AAZ28826  
ID AAZ28826 standard; DNA; 19 BP.

XX AC AAZ28826;

XX DT 01-FEB-2000 (first entry)

XX DE Rat membrane metalloprotease NEPII gene probe #16.

XX KW Rat; membrane metalloprotease; neprilysine II; NEPII; inactivation; ss;  
neuron; hormone; peptide messenger; inhibitor; detection; disorder;  
cardiovascular disease; neurodegenerative disease; growth disorder;  
hypothalamic-hypophyseal axis; endocrine disorder; probe; hybridisation.

XX OS Synthetic.

XX OS Rattus rattus.

XX PN FR2777291-A1.

XX PD 15-OCT-1999.

XX PF 08-APR-1998; 98FR-0004389.

XX PR 08-APR-1998; 98FR-0004389.

XX PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

XX PI Ouimet T, Gros C, Haret C, Bonhomme MC, Facchinetti P;

XX PI Schwartz JC;

XX DR WPI; 1999-593429/51.



```
XX
SQ Sequence 2765 BP; 684 A; 735 C; 787 G; 559 T; 0 other;
Query Match 100.0%; Score 19; DB 20; Length 2765;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTTCGCCATCCAGTCCATC 19
    |||||||
DB 2281 GTTCGCCATCCAGTCCATC 2299

RESULT 4
ABV14605
ID ABV14605 standard; cDNA; 324 BP.
AC ABV14605;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 14596.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
WPI; 2001-662795/76.
XX
Novel isolated nucleic acid molecule associated with cancerous state of
prostate cells and correlating with presence of prostate cancer, useful
for detecting presence of prostate cancer, stage of prostate cancer -
Claim 1; Page 2438; 11750pp; English.
XX
The invention relates to an isolated nucleic acid molecule (I) comprising
a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
specification or its complement. (I) is useful for:
(a) assessing whether a patient is afflicted with prostate cancer;
(b) monitoring the progression of prostate cancer in a patient;
(c) assessing the efficacy of a test compound to inhibit prostate
cancer in a patient;
(d) assessing the efficacy of a therapy for inhibiting prostate cancer
in a patient;
(e) selecting a composition for inhibiting prostate cancer in a patient;
(f) assessing the prostate cell carcinogenic potential of a compound;
(g) determining whether prostate cancer has metastasized in a patient;
(h) assessing the aggressiveness or indolence of prostate cancer in a
patient;
(i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
Sequence 324 BP; 70 A; 101 C; 98 G; 55 T; 0 other;
Query Match 91.6%; Score 17.4; DB 23; Length 324;
Best Local Similarity 94.7%; Pred. No. 47;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GTTCGCCATCCAGTCCATC 19
    |||||||
DB 290 GTTCGCCATCCAGTCCATC 308

RESULT 5
ABV05436
ID ABV05436 standard; cDNA; 340 BP.
AC ABV05436;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 5427.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
WPI; 2001-662795/76.
XX
Novel isolated nucleic acid molecule associated with cancerous state of
prostate cells and correlating with presence of prostate cancer, useful
for detecting presence of prostate cancer, stage of prostate cancer -
Claim 1; Page 915; 11750pp; English.
XX
The invention relates to an isolated nucleic acid molecule (I) comprising
a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
specification or its complement. (I) is useful for:
(a) assessing whether a patient is afflicted with prostate cancer;
(b) monitoring the progression of prostate cancer in a patient;
(c) assessing the efficacy of a test compound to inhibit prostate
cancer in a patient;
(d) assessing the efficacy of a therapy for inhibiting prostate cancer
in a patient;
(e) selecting a composition for inhibiting prostate cancer in a patient;
(f) assessing the prostate cell carcinogenic potential of a compound;
(g) determining whether prostate cancer has metastasized in a patient;
(h) assessing the aggressiveness or indolence of prostate cancer in a
patient;
(i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
Sequence 340 BP; 69 A; 109 C; 106 G; 56 T; 0 other;
Query Match 91.6%; Score 17.4; DB 23; Length 340;
Best Local Similarity 94.7%; Pred. No. 47;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GTTCGCCATCCAGTCCATC 19
    |||||||
DB 290 GTTCGCCATCCAGTCCATC 308
```

## RESULT 6

ABV35683  
ID ABV35683 standard; cDNA; 569 BP.  
XX  
AC ABV35683;  
XX  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 35674.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -

Claim 1; Page 7417; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:  
(a) assessing whether a patient is afflicted with prostate cancer;  
(b) monitoring the progression of prostate cancer in a patient;  
(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;  
(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;  
(e) selecting a composition for inhibiting prostate cancer in a patient;  
(f) assessing the prostate cell carcinogenic potential of a compound;  
(g) determining whether prostate cancer has metastasized in a patient;  
(h) assessing the aggressiveness or indolence of prostate cancer in a patient;  
(I) is also useful as a pharmacodynamic or pharmacogenomic marker.

Sequence 569 BP; 121 A; 166 C; 165 G; 116 T; 1 other;

Query Match 91.6%; Score 17.4; DB 23; Length 569;  
Best Local Similarity 94.7%; Pred. No. 49;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19  
|||||  
DB 322 GTTCGCCATCCATCCATC 340

## RESULT 7

ABV44488  
ID ABV44488 standard; cDNA; 569 BP.  
XX  
AC ABV44488;  
XX

DT 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 44479.

XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

XX  
PN WO200160860-A2.

XX  
PD 23-AUG-2001.

XX  
PF 20-FEB-2001; 2001WO-US05171.

XX  
PR 17-FEB-2000; 2000US-183319P.

XX  
PR 16-MAR-2000; 2000US-189862P.

XX  
PR 25-MAY-2000; 2000US-207454P.

XX  
PR 09-JUN-2000; 2000US-211314P.

XX  
PR 18-JUL-2000; 2000US-219007P.

XX  
PR 13-DEC-2000; 2000US-255281P.

XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX  
PI Schlegel R, Endege WO, Monahan JE;

XX  
DR WPI; 2001-662795/76.

XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -

XX  
PS Claim 1; Page 8827; 11750pp; English.

XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:

XX  
CC (a) assessing whether a patient is afflicted with prostate cancer;

XX  
CC (b) monitoring the progression of prostate cancer in a patient;

XX  
CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;

XX  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;

XX  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;

XX  
CC (f) assessing the prostate cell carcinogenic potential of a compound;

XX  
CC (g) determining whether prostate cancer has metastasized in a patient;

XX  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;

XX  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX  
SQ Sequence 569 BP; 121 A; 166 C; 165 G; 116 T; 1 other;

Query Match 91.6%; Score 17.4; DB 23; Length 569;

Best Local Similarity 94.7%; Pred. No. 49;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19

|||||

DB 322 GTTCGCCATCCATCCATC 340

## RESULT 8

AAH26248  
ID AAH26248 standard; cDNA; 1124 BP.  
XX  
AC AAH26248;  
XX  
DT 02-OCT-2001 (first entry)  
XX  
DE Human secreted protein cDNA clone HSIDD62.

XX  
KW Secreted protein; human; HSIDD62; colon cancer; ovary cancer;  
KW tumour; diagnosis; gene therapy; ss.

XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 XX CDS 2..367  
 XX FT /\*tag="a  
 XX FT /partial  
 XX PN WO200153343-A1.  
 XX XX  
 XX PD 26-JUL-2001.  
 XX XX  
 XX PF 17-JAN-2001; 2001WO-US01436.  
 XX XX  
 XX PR 18-JAN-2000; 2000US-0176307.  
 XX XX  
 XX PA (HUMA-) HUMAN GENOME SCI INC.  
 XX PI Ruben SM, Shi Y;  
 XX XX  
 XX DR WPI; 2001-483137/52.  
 XX DR P-PSDB; AAB82596.  
 XX XX  
 XX PT Nucleic acid molecules encoding human secreted proteins, used in  
 XX PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's  
 XX PT diseases and cancers  
 XX PS  
 XX PS Claim 1(a): Page 330; 339pp; English.  
 XX XX  
 XX CC The present sequence is that of cDNA clone HSIDD62 (deposited as  
 XX CC ATCC PTA1201), which encodes a novel human secreted protein (see  
 XX CC AAB82596) that shows homology to atrial natriuretic protein  
 XX CC binding protein (ANBPB), which is thought to have diuretic  
 XX CC (particularly natriuretic) and hypotensive activities. The gene  
 XX CC is expressed in adult small intestine tissue as well as in colon  
 XX CC tumour tissue and in ovarian cancer tissue. This tissue  
 XX CC distribution and homology to human ANBPB indicate that  
 XX CC polynucleotides, translation products and antibodies corresponding  
 XX CC to this gene could be useful in the diagnosis, detection and/or  
 XX CC treatment of diseases and/or disorders of the gastrointestinal  
 XX CC system, as well as colon tumours and ovarian cancer. Antagonists  
 XX CC directed against the gene translation product may be useful in  
 XX CC prohibiting or eliminating its biological activities, such as  
 XX CC diuretic and hypotensive activities. Vectors, host cells,  
 XX CC antibodies, recombinant methods for producing the secreted protein,  
 XX CC and screening methods are also provided.  
 XX XX  
 XX SQ Sequence 1124 BP; 251 A; 362 C; 320 G; 191 T; 0 other;  
 XX XX  
 XX Query Match 91.6%; Score 17.4; DB 22; Length 1124;  
 XX Best Local Similarity 94.7%; Pred. No. 52;  
 XX Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 XX XX  
 XX QY 1 GTTCGCCATCCAGTCCATC 19  
 XX DB ||||||||||| |||||  
 XX 608 GTTCGCCATCCATCCATC 626  
 XX XX  
 XX RESULT 9  
 XX AAH35007  
 XX ID AAH35007 standard; cDNA; 1226 BP.  
 XX XX  
 XX AC AAH35007;  
 XX XX  
 XX DT 03-SEP-2001 (first entry)  
 XX XX  
 XX DE Human colon cancer antigen encoding cDNA SEQ ID NO:2089.  
 XX XX  
 XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
 XX KW colorectal carcinoma; ss.  
 XX XX  
 XX OS Homo sapiens.  
 XX XX

PN WO200122920-A2.  
 XX XX  
 XX PD 05-APR-2001.  
 XX XX  
 XX PF 28-SEP-2000; 2000WO-US26524.  
 XX XX  
 XX PR 29-SEP-1999; 99US-0157137.  
 XX PR 03-NOV-1999; 99US-0163280.  
 XX XX  
 XX PA (HUMA-) HUMAN GENOME SCI INC.  
 XX XX  
 XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
 XX XX  
 XX DR WPI; 2001-235357/24.  
 XX DR P-PSDB; AAG75602.  
 XX XX  
 XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
 XX PT useful for preventing, diagnosing and/or treating colorectal cancers -  
 XX XX  
 XX PS Claim 1; Page 3574; 9803pp; English.  
 XX XX  
 XX CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
 XX CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
 XX CC the proteins are collectively known as colon cancer antigens. The colon  
 XX CC cancer antigens have cytostatic activity and can be used in gene  
 XX CC therapy and vaccine production. N and P may be used in the prevention,  
 XX CC diagnosis and treatment of diseases associated with inappropriate P  
 XX CC expression. For example, N and P may be used to treat disorders  
 XX CC associated with decreased expression by rectifying mutations or deletions  
 XX CC in a patient's genome that affect the activity of P by expressing  
 XX CC inactive proteins or to supplement the patient's own production of P.  
 XX CC Additionally, N may be used to produce the colon cancer-associated Ps  
 XX CC by inserting the nucleic acids into a host cell and culturing the cell  
 XX CC to express the proteins. N and P can be used in the prevention, diagnosis  
 XX CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
 XX CC and AAB77789 represent sequences used in the exemplification of the  
 XX CC present invention.  
 XX CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
 XX CC missing at time of publication, meaning no sequences are present for  
 XX CC SEQ ID NO:1027 to 1052, 7921 and 7922.  
 XX XX  
 XX SQ Sequence 1226 BP; 270 A; 388 C; 345 G; 213 T; 10 other;  
 XX XX  
 XX Query Match 91.6%; Score 17.4; DB 22; Length 1226;  
 XX Best Local Similarity 94.7%; Pred. No. 53;  
 XX Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 XX XX  
 XX QY 1 GTTCGCCATCCAGTCCATC 19  
 XX DB ||||||||||| |||||  
 XX 606 GTTCGCCATCCATCCATC 624  
 XX XX  
 XX RESULT 10  
 XX ABV23433/C  
 XX ID ABV23433 standard; cDNA; 1746 BP.  
 XX XX  
 XX AC ABV23433;  
 XX XX  
 XX DT 16-SEP-2002 (first entry)  
 XX XX  
 XX DE Human prostate expression marker cDNA 23424.  
 XX XX  
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 XX KW pharmacogenomic marker; gene; ss.  
 XX XX  
 XX OS Homo sapiens.  
 XX XX  
 XX PN WO200160860-A2.  
 XX XX  
 XX PD 23-AUG-2001.  
 XX XX  
 XX PF 20-FEB-2001; 2001WO-US05171.  
 XX XX



PN WO200136610-A1.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-BP11532.

XX 19-NOV-1999; 99EP-0203862.

PR 19-NOV-1999; 99NL-1013616.

PR 31-MAY-2000; 2000EP-0201937.

PR 31-MAY-2000; 2000NL-1015356.

XX (SOLV ) SOLVAY PHARM BV.

XX Deleersnijder W, Wiegers R, Weske M;

WPI; 2001-343815/36.

DR P-PSDB; AAB83840.

XX New IGS5 polypeptides useful for treating infections, pain, cancer,  
PT diabetes, obesity, anorexia, bulimia, asthma, schizophrenia,  
PT hypertension, urinary retention and Parkinson's disease

XX Claim 11; Page 5-6; 115pp; English.

XX The present sequence encodes a human metalloprotease enzyme designated  
CC IGS5. IGS5 polynucleotides and polypeptides are useful for treating  
CC infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma,  
CC Parkinson's disease, acute heart failure, hypotension, hypertension,  
CC urinary retention, osteoporosis, angina pectoris, myocardial infarction,  
CC stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,  
CC psychotic and neurological disorders, autism, multiple sclerosis,  
CC Alzheimer's disease, and other neurodegenerative diseases, sleep  
CC disorder, epilepsy, cardiovascular diseases, arteriosclerosis,  
CC cerebrovasospasm, subarachnoid hemorrhage, cerebral ischemia, cerebral  
CC infarction, peripheral vascular disease, Raynaud's disease, kidney  
CC diseases, gastrointestinal disorders, motility disorders and conditions  
CC of delayed gastric emptying, post-operative or diabetic gastroparesis,  
CC diarrhoea, inflammations, chemotherapy induced injury, tumour invasion,  
CC immune disorders, arthritis, endotoxin shock, sepsis, complications of  
CC diabetes mellitus, and severe mental retardation and dyskinesias, such  
CC as Huntington's disease or Gilles de la Tourette's syndrome.

XX Sequence 2076 BP; 493 A; 578 C; 636 G; 369 T; 0 other;

Query Match 91.6%; Score 17.4; DB 22; Length 2076;

Best Local Similarity 94.7%; Pred. No. 55;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19

DB 1926 GTTCGCCATCCATCCATC 1944

RESULT 13

AAS97186

ID AAS97186 standard; cDNA; 2232 BP.

XX AAS97186;

DT 26-FEB-2002 (first entry)

DE Human metalloprotease partial DNA sequence #15.

XX Human; protease; PCR primer; cytostatic; immunomodulator; cardiant;  
KW vasotropic; antimigraine; analgesic; endocrine; nootropic; tranquiliser;  
KW hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic;  
KW anorectic; antiinflammatory; aspartyl protease; cysteine protease;  
KW metalloprotease; serine protease; cancer; haematopoietic; breast;  
KW lung; prostate; cervical; brain; ovarian; bladder; kidney; pain;  
KW immune-related disease; cardiovascular disease; neuronal disease;  
KW migraine; sexual dysfunction; mood disorder; attention disorder;  
KW cognition disorder; hypotension; hypertension; psychotic disorder;  
KW dyskinesia; metabolic disorder; inflammatory disorder; ss.

XX Homo sapiens.

XX WO200183782-A2.

XX 08-NOV-2001.

XX 04-MAY-2001; 2001WO-US14431.

XX 04-MAY-2000; 2000US-201879P.

XX (SUGF-) SUGEN INC.

XX Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;

PI Payne V;

XX WPI; 2002-041502/05.

DR P-PSDB; AAU72903.

XX Novel protease polypeptide useful for screening for substances that may  
PT be used to treat, e.g., cancers, immune-related diseases,  
PT cardiovascular disease, migraine, pain, psychotic and inflammatory  
PT disorders

XX Claim 30; Figure 1R-S; 232pp; English.

XX The invention relates to an isolated, enriched, or purified protease  
CC polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to  
CC screen for substances (S) that may modulate its activity. Administering  
CC S (which modulates protease activity in vitro) may be used to treat a  
CC disease or disorder selected from cancers (e.g., of tissues of blood or  
CC haematopoietic origin, of the breast, colon, lung, prostate, cervical,  
CC brain, ovarian, bladder or kidney), immune-related diseases and  
CC disorders; cardiovascular disease, brain or neuronal-associated diseases  
CC (e.g., central or peripheral nervous system diseases, migraine, pain,  
CC sexual dysfunction, mood disorders, attention disorders, cognition  
CC disorders, hypotension, hypertension, psychotic disorders, neurological  
CC disorders and dyskinesias), metabolic disorders and inflammatory  
CC disorders. (I) may also be useful as a diagnostic tool for a disease or  
CC disorder such as those above. AAS97159-AAS97195 represent human  
CC protease coding sequences and primers of the invention.

XX Sequence 2232 BP; 512 A; 620 C; 705 G; 395 T; 0 other;

Query Match 91.6%; Score 17.4; DB 24; Length 2232;

Best Local Similarity 94.7%; Pred. No. 56;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19

DB 2082 GTTCGCCATCCATCCATC 2100

RESULT 14

AAS9739

ID AAS9739 standard; DNA; 2262 BP.

XX AAS9739;

DT 23-JUL-2001 (first entry)

DE Nucleotide sequence of a human metalloprotease enzyme IGS5.

XX Metalloprotease; IGS5; infection; pain; cancer; diabetes; obesity;  
KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;  
KW hypotension; hypertension; urinary retention; osteoporosis;  
KW angina pectoris; myocardial infarction; stroke; ulcer; allergy;  
KW benign prostatic hypertrophy; migraine; psychotic disorder;  
KW neurological disorder; autism; multiple sclerosis; Alzheimer's disease;  
KW neurodegenerative disease; sleep disorder; epilepsy; kidney disease;  
KW cardiovascular disease; arteriosclerosis; cerebrovasospasm;  
KW subarachnoid hemorrhage; cerebral ischemia; cerebral infarction;  
KW peripheral vascular disease; Raynaud's disease; motility disorder;



gastrointestinal disorder; gastric emptying; gastroparesis; diarrhoea; inflammation; chemotherapy induced injury; tumour invasion; immune disorder; arthritis; endotoxin shock; sepsis; diabetes mellitus; severe mental retardation; dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 1..2262

/\*tag= a

/product= "metalloprotease enzyme IGS5"

WO200136610-A1.

25-MAY-2001.

17-NOV-2000; 2000WO-EP11532.

19-NOV-1999; 99EP-0203862.

19-NOV-1999; 99NL-1013616.

31-MAY-2000; 2000EP-0201937.

31-MAY-2000; 2000NL-1015356.

(SOLV ) SOLVAY PHARM BV.

Deleersnijder W, Wiegers R, Weske M;

WPT; 2001-343815/36.

P-PSDB; AAB83842.

New IGS5 polypeptides useful for treating infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma, schizophrenia, hypertension, urinary retention and Parkinson's disease

Claim 11; Page 8-9; 115pp; English.

The present sequence encodes a human metalloprotease enzyme designated IGS5. IGS5 polynucleotides and polypeptides are useful for treating infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, angina pectoris, myocardial infarction, stroke, ulcers, allergies, benign prostatic hypertrophy, migraine, psychotic and neurological disorders, autism, multiple sclerosis, Alzheimer's disease, and other neurodegenerative diseases, sleep disorder, epilepsy, cardiovascular diseases, arteriosclerosis, cerebrovasospasm, subarachnoid hemorrhage, cerebral ischemia, cerebral infarction, peripheral vascular disease, Raynaud's disease, kidney diseases, gastrointestinal disorders, motility disorders and conditions of delayed gastric emptying, post-operative or diabetic gastroparesis, diarrhoea, inflammations, chemotherapy induced injury, tumour invasion, immune disorders, arthritis, endotoxin shock, sepsis, complications of diabetes mellitus, and severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome.

Sequence 2262 BP; 520 A; 628 C; 716 G; 398 T; 0 other;

Query Match 91.6%; Score 17.4; DB 22; Length 2262;

Best Local Similarity 94.7%; Pred. No. 56;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GTTCCCATCCAGTCCATC 19

||||| ||||| |||||

Db 2112 GTTCCCATCCATCCATC 2130

RESULT 15

RA030580

ID A030580 standard; cDNA; 2318 BP.

XX AC

AA030580;

XX XX

DT 21-MAY-2002 (first entry)

XX

DE

Human protease, PRYS-13 cDNA.

XX

Human; protease; PRYS-13; enzyme; gastritis; cirrhosis; Crohn's disease;

gastrointestinal disorder; autoimmune; inflammatory; cell proliferative;

cardiovascular; developmental; epithelial; neurological; reproductive;

AIDS; Acquired Immune Deficiency Syndrome; allergy; rheumatoid arthritis;

anaemia; asthma; atherosclerosis; hypertension; myocardial infarction;

hepatitis; cancer; psoriasis; Cushing's syndrome; hypothyroidism; eczema;

epilepsy; Alzheimer's disease; Huntington's disease; Parkinson's disease;

Pick's disease; infertility; vitiligo; drug screening; gene therapy; ss.

XX

Homo sapiens.

XX

Key Location/Qualifiers

CDS 338..1651

/\*tag= a

/product= "Human PRYS-13 protein"

sig\_peptide 338..427

/\*tag= b

mat\_peptide 428..1648

/\*tag= c

/product= "Mature PRYS-13 protein"

XX

WO200208396-A2.

XX

31-JAN-2002.

XX

17-JUL-2001; 2001WO-US22397.

XX

21-JUL-2000; 2000US-220063P.

XX

28-JUL-2000; 2000US-221680P.

XX

04-AUG-2000; 2000US-223544P.

XX

11-AUG-2000; 2000US-224717P.

XX

16-AUG-2000; 2000US-225988P.

XX

23-AUG-2000; 2000US-227568P.

XX

(INCY-) INCYTE GENOMICS INC.

XX

Deleese AM, Gandhi AR, Hafalia AJA, Lu DAM, Patterson C;

Tribouley CM, Das D, Kallick DA, Nguyen DB, Lee EA, Khan FA;

Yue H, Au-Young J, Griffin JA, Policky JL, Ramkumar J, Yang J;

Thangavelu K, Ding L, Kearney L, Baughn MR, Borowsky M;

Sanjanwala MS, Yao MG, Burford N, Wallia NK, Lal P, Lee S, Todd S;

Lo TP, Tang YT, Elliott VS, Azimzai Y, Lu Y;

XX

WPI; 2002-206082/26.

XX

P-PSDB; AAE19176.

XX

New human protease polypeptide, useful in diagnosis, prevention and

treatment of gastrointestinal, cardiovascular, autoimmune/inflammatory,

cell proliferative, developmental, epithelial and neurological

disorders

XX

Claim 5; Page 174-175; 182pp; English.

XX

The invention relates to an isolated human protease polypeptide (PRYS).

XX

PRYS protein and DNA are useful for diagnosing, treating and preventing

gastrointestinal disorders (gastritis, cirrhosis, Crohn's disease),

autoimmune/inflammatory disorders (AIDS, allergy, rheumatoid arthritis,

anaemia, asthma), cardiovascular disorder (atherosclerosis, hypertension,

myocardial infarction), cell proliferative disorders (hepatitis, cancer,

psoriasis), developmental disorders (Cushing's syndrome, hypothyroidism),

epithelial disorder (vitiligo, keloid, eczema), neurological disorders

(epilepsy, Alzheimer's disease, Pick's disease, Huntington's disease,

Parkinson's disease), and reproductive disorders (infertility). PRYS

protein is useful in a number of drug screening techniques and to

analyse the proteome of a tissue or cell type. PRYS DNA is useful for

creating knockin humanised animals or transgenic animals to model human

diseases, in somatic or germline gene therapy and in microarrays

utilising fluids or tissues from patients to detect altered PKIN

expression. The present sequence is human PRYS-13 cDNA.

SQ Sequence 2318 BP; 526 A; 647 C; 735 G; 410 T; 0 other;

Query Match 91.6%; Score 17.4; DB 24; Length 2318;  
Best Local Similarity 94.7%; Pred. No. 56;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19

|||||

Db 2168 GTTCGCCATCCCAATCCATC 2186

Search completed: July 8, 2003, 02:19:14  
Job time : 120.376 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 01:24:03 ; Search time 26.2293 Seconds  
(without alignments)  
222.151 Million cell updates/sec

Title: US-09-647-780A-20

Perfect score: 19

Sequence: 1 gttcgccatccagtcctc 19

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents\_NA:  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	15.8	83.2	449	US-09-615-192A-248	Sequence 248, Appl
2	15.8	83.2	2446	US-09-310-463-11	Sequence 11, Appl
3	15.8	83.2	2446	US-08-842-248A-11	Sequence 11, Appl
4	15.4	81.1	2055	US-09-134-078-58	Sequence 58, Appl
5	15.4	81.1	2280	US-09-346-408-3	Sequence 3, Appl
6	15.8	78.9	1409	US-08-928-213B-58	Sequence 58, Appl
7	14.8	77.9	1174	US-08-459-586-17	Sequence 17, Appl
8	14.8	77.9	1174	US-08-282-696-17	Sequence 17, Appl
9	14.8	77.9	1328	US-08-459-586-11	Sequence 11, Appl
10	14.8	77.9	1328	US-08-282-696-11	Sequence 11, Appl
11	14.8	77.9	2582	US-08-816-105A-2	Sequence 11, Appl
12	14.8	77.9	4880	US-09-031-563-1	Sequence 2, Appl
13	14.8	77.9	4880	US-09-392-277-1	Sequence 1, Appl
14	14.8	77.9	5125	US-09-031-563-4	Sequence 1, Appl
15	14.8	77.9	5125	US-09-392-277-4	Sequence 4, Appl
16	14.8	77.9	5352	US-09-031-563-6	Sequence 4, Appl
17	14.8	77.9	5252	US-09-392-277-6	Sequence 6, Appl
18	14.8	77.9	18994	US-08-459-586-4	Sequence 4, Appl
19	14.8	77.9	18994	US-08-282-696-4	Sequence 4, Appl
20	14.4	75.8	510	US-09-164-193-3	Sequence 3, Appl
21	14.4	75.8	510	US-09-221-448A-3	Sequence 3, Appl
22	14.4	75.8	1315	US-09-164-193-1	Sequence 1, Appl
23	14.4	75.8	1315	US-09-221-448A-1	Sequence 1, Appl
24	14.4	75.8	1664	US-09-339-993-1	Sequence 1, Appl
25	14.4	75.8	2575	US-09-077-354B-1	Sequence 1, Appl
26	14.4	75.8	2925	US-08-462-484-9	Sequence 9, Appl
27	14.4	75.8	2925	US-08-441-147-9	Sequence 9, Appl

28 14.4 75.8 2925 5 PCT-US95-07536-9 Sequence 9, Appl  
29 14.4 75.8 3506 3 US-09-091-893-9 Sequence 9, Appl  
30 14.4 75.8 4285 4 US-09-040-774-1 Sequence 1, Appl  
31 14.4 75.8 5410 4 US-09-221-017B-70 Sequence 70, Appl  
32 14.4 75.8 6048 4 US-09-634-920-3 Sequence 3, Appl  
33 14.4 75.8 10380 4 US-09-077-354B-3 Sequence 3, Appl  
34 14.2 74.7 532 3 US-08-824-692-26 Sequence 26, Appl  
35 14.2 74.7 590 3 US-08-824-692-28 Sequence 28, Appl  
36 14.2 74.7 615 1 US-08-247-946A-5 Sequence 5, Appl  
37 14.2 74.7 615 5 PCT-US95-06420-5 Sequence 5, Appl  
38 14.2 74.7 688 3 US-08-824-692-27 Sequence 27, Appl  
39 14.2 74.7 767 3 US-08-824-692-25 Sequence 25, Appl  
40 14.2 74.7 1455 4 US-09-134-001C-2565 Sequence 2565, Ap  
41 14.2 74.7 1605 4 US-09-310-463-7 Sequence 7, Appl  
42 14.2 74.7 1605 4 US-08-842-248A-7 Sequence 7, Appl  
43 14.2 74.7 1620 3 US-08-985-950-11 Sequence 11, Appl  
44 14.2 74.7 1725 4 US-09-310-463-15 Sequence 15, Appl  
45 14.2 74.7 1725 4 US-08-842-248A-15 Sequence 15, Appl

## ALIGNMENTS

RESULT 1  
US-09-615-192A-248/c  
; Sequence 248, Application US/09615192A  
; Patent No. 6410718  
; GENERAL INFORMATION:  
; APPLICANT: Bloksberg, Leonard N.  
; APPLICANT: Havukkala, Ilkka  
; TITLE OF INVENTION: Materials and Methods for the  
; TITLE OF INVENTION: Modification of Plant Lignin Content  
; FILE REFERENCE: 11000.100364U  
; CURRENT APPLICATION NUMBER: US/09/615,192A  
; CURRENT FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 08/975,316  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: US 08/713,000  
; PRIOR FILING DATE: 1996-09-11  
; PRIOR APPLICATION NUMBER: US 09/169,789  
; PRIOR FILING DATE: 1998-10-09  
; NUMBER OF SEQ ID NOS: 405  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 248  
; LENGTH: 449  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-09-615-192A-248

Query Match 83.2% ; Score 15.8; DB 4; Length 449;  
Best Local Similarity 89.5% ; Pred. No. 57;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GTTCGCCATCCAGTCCTC 19  
|||||||  
Db 415 GTTCGCCATCCAGTCCTC 397

RESULT 2  
US-09-310-463-11  
; Sequence 11, Application US/09310463A  
; Patent No. 6384203  
; GENERAL INFORMATION:  
; APPLICANT: Cosman, David J.  
; APPLICANT: Anderson, Dirk M.

; APPLICANT: Borges, Luis  
; TITLE OF INVENTION: Family of Immunoregulators Designated Leukocyte Immunoglobul  
; TITLE OF INVENTION: Like Receptors (LIR)  
; FILE REFERENCE: 2624-A  
; CURRENT APPLICATION NUMBER: US/09/310,463A  
; CURRENT FILING DATE: 1999-05-12  
; EARLIER APPLICATION NUMBER: 08/842,248  
; EARLIER FILING DATE: 1997-04-24

NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 2446  
; TYPE: DNA  
; ORGANISM: human  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (171)..(1040)  
US-09-310-463-11

Query Match 83.2%; Score 15.8; DB 4; Length 2446;  
Best Local Similarity 89.5%; Pred. No. 67;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTCCCATCCAGTCCATC 19  
||||| ||||| ||||| ||||| |||||

Db 416 GTTCCCATCCAGTCCATC 434

## RESULT 3

US-08-842-248A-11  
; Sequence 11, Application US/08842248A  
; Patent No. 6448035  
; GENERAL INFORMATION:  
; APPLICANT: Cosman, David J.  
; TITLE OF INVENTION: Family of Immunoregulators Designated  
; TITLE OF INVENTION: Leukocyte Immunoglobulin-Like Receptors (LIR)  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Janis C. Henry, Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: US  
; ZIP: 98101

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM/PC Compatible  
; OPERATING SYSTEM: Microsoft Word 7.0  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/842,248A  
; FILING DATE: April 24, 1997

CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Henry, Janis C.  
; REGISTRATION NUMBER: 34,347  
; REFERENCE/DOCKET NUMBER: 2624  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; TELEX: 756822

INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2446 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; IMMEDIATE SOURCE:  
; CLONE: pbm36-2  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 171..1037  
US-08-842-248A-11

Query Match 83.2%; Score 15.8; DB 4; Length 2446;  
Best Local Similarity 89.5%; Pred. No. 67;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTCCCATCCAGTCCATC 19  
||||| ||||| ||||| ||||| |||||

Db 416 GTTCCCATCCAGTCCATC 434

## RESULT 4

US-09-134-078-58  
; Sequence 58, Application US/09134078  
; Patent No. 6368844  
; GENERAL INFORMATION:  
; APPLICANT: Bylina, Edward J.  
; TITLE OF INVENTION: GLYCOSIDASE ENZYMES  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Gray Cary Ware & Freidenrich LLP  
; STREET: 4365 Executive Drive, Suite 1600  
; CITY: San Diego  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92121

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/134,078  
; FILING DATE: 13-AUG-1998  
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/949,026  
; FILING DATE: 10-OCT-1997  
; APPLICATION NUMBER: 60/056,916  
; FILING DATE: 06-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halle, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 09010/024002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 858/677-1456  
; TELEFAX: 858/677-1465

INFORMATION FOR SEQ ID NO: 58:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2035 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 1...2052  
US-09-134-078-58

Query Match 81.1%; Score 15.4; DB 4; Length 2055;  
Best Local Similarity 94.1%; Pred. No. 1e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTCGCCATCCAGTCCAT 18  
||||| ||||| ||||| ||||| |||||

Db 1380 TTCGCCATCCAGTCCAT 1396

## RESULT 5

US-09-346-408-3  
; Sequence 3, Application US/09346408B  
; Patent No. 6338966  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Anderson, Shawn  
; APPLICANT: Falco, Carl  
; APPLICANT: Rafalski, Antoni  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-A  
; CURRENT APPLICATION NUMBER: US/09/346,408B  
; CURRENT FILING DATE: 1999-07-01

EARLIER APPLICATION NUMBER: 60/092,833  
EARLIER FILING DATE: July 14, 1998  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 3  
LENGTH: 2280  
TYPE: DNA  
ORGANISM: Momordica charantia  
US-09-346-408-3

Query Match 81.1%; Score 15.4; DB 4; Length 2280;  
Best Local Similarity 94.1%; Pred. No. 1e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TTCGCCATCCAGTCCAT 18  
||||| |||||||  
Db 825 TTCGCATCCAGTCCAT 841

## RESULT 6

US-08-928-213B-58  
Sequence 58, Application US/08928213B  
Patent No. 6238905

## GENERAL INFORMATION:

APPLICANT: McHenry, Charles S.  
Seville, Mark  
Cull, Millard G.

TITLE OF INVENTION: NOVEL THERMOPHILIC POLYMERASE III  
HOLOENZYME

NUMBER OF SEQUENCES: 195

CORRESPONDENCE ADDRESS:

ADDRESSEE: MEDLEN & CARROLL, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/928,213B

FILING DATE: 12-Sep-1997

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: MacKnight, Kamrin T.

REGISTRATION NUMBER: 38,230

REFERENCE/DOCKET NUMBER: ENZYCO-02550

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-705-8410

TELEFAX: 415-397-8338

SEQUENCE CHARACTERISTICS:

LENGTH: 1109 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "DNA"

SEQUENCE DESCRIPTION: SEQ ID NO: 58:

US-08-928-213B-58

Query Match 78.9%; Score 15; DB 4; Length 1109;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TCGCCATCCAGTCCA 17  
||||| |||||||

Db 687 TCGCCATCCAGTCCA 701

## RESULT 7

US-08-459-586-17  
Sequence 17, Application US/08459586

Patent No. 5720957

GENERAL INFORMATION:

APPLICANT: Jones, Thomas R.

ADDRESSEE: Campbell, Ann E.

TITLE OF INVENTION: Identification of a Human

Cytomegalovirus Gene Region Involved in Down-Regulation of

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: American Cyanamid Company

STREET: One Campus Drive

CITY: Parsippany

STATE: New Jersey

COUNTRY: United States

ZIP: 07054

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/459,586

FILING DATE: 02-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Barnhard, Elizabeth M.

REGISTRATION NUMBER: 31,088

REFERENCE/DOCKET NUMBER: 32,425-01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-683-2158

TELEFAX: 201-683-4117

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 1174 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-459-586-17

Query Match 77.9%; Score 14.8; DB 1; Length 1174;

Best Local Similarity 88.9%; Pred. No. 1.8e+02;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GTTCGCCATCCAGTCCAT 18  
||||| |||||||

Db 845 GTTCGCCATCCAGTCCAT 862

## RESULT 8

US-08-282-696-17

Sequence 17, Application US/08282696

Patent No. 5846806

GENERAL INFORMATION:

APPLICANT: Jones, Thomas R.

ADDRESSEE: Campbell, Ann E.

TITLE OF INVENTION: Identification of a Human

Cytomegalovirus Gene Region Involved in Down-Regulation of

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: American Cyanamid Company

STREET: One Campus Drive

CITY: Parsippany

STATE: New Jersey

COUNTRY: United States

ZIP: 07470-8426

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/282,696  
FILING DATE: 29-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Barnhard, Elizabeth M.  
REGISTRATION NUMBER: 31,088  
REFERENCE/DOCKET NUMBER: 32,425-00  
TELEPHONE: 201-831-3246  
TELEFAX: 201-831-3305  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1174 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-282-696-17

Query Match. 77.9%; Score 14.8; DB 2; Length 1174;  
Best Local Similarity 88.9%; Pred. No. 1.8e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCAT 18  
Db 845 GTGCGCTTCCAGTCCAT 862

RESULT 9  
US-08-459-586-11  
Sequence 11, Application US/08459586  
Patent No. 5720957  
GENERAL INFORMATION:  
APPLICANT: Jones, Thomas R.  
APPLICANT: Campbell, Ann E.  
TITLE OF INVENTION: Identification of a Human  
TITLE OF INVENTION: Cytomegalovirus Gene Region Involved in Down-Regulation of  
TITLE OF INVENTION: MHC Class I Heavy Chain Expression  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: American Cyanamid Company  
STREET: One Campus Drive  
CITY: Parsippany  
STATE: New Jersey  
COUNTRY: United States  
ZIP: 07054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,586  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Barnhard, Elizabeth M.  
REGISTRATION NUMBER: 31,088  
REFERENCE/DOCKET NUMBER: 32,425-01  
TELEPHONE: 201-683-2158  
TELEFAX: 201-683-4117  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1328 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-459-586-11

Query Match. 77.9%; Score 14.8; DB 1; Length 1328;  
Best Local Similarity 88.9%; Pred. No. 1.9e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCAT 18  
Db 642 GTGCGCTTCCAGTCCAT 659

RESULT 10  
US-08-282-696-11  
Sequence 11, Application US/08282696  
Patent No. 5845806  
GENERAL INFORMATION:  
APPLICANT: Jones, Thomas R.  
APPLICANT: Campbell, Ann E.  
TITLE OF INVENTION: Identification of a Human  
TITLE OF INVENTION: Cytomegalovirus Gene Region Involved in Down-Regulation of  
TITLE OF INVENTION: MHC Class I Heavy Chain Expression  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: American Cyanamid Company  
STREET: One Cyanamid Plaza  
CITY: Wayne  
STATE: New Jersey  
COUNTRY: United States  
ZIP: 07470-8426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/282,696  
FILING DATE: 29-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Barnhard, Elizabeth M.  
REGISTRATION NUMBER: 31,088  
REFERENCE/DOCKET NUMBER: 32,425-00  
TELEPHONE: 201-831-3246  
TELEFAX: 201-831-3305  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1328 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-282-696-11

Query Match. 77.9%; Score 14.8; DB 2; Length 1328;  
Best Local Similarity 88.9%; Pred. No. 1.9e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCAT 18  
Db 642 GTGCGCTTCCAGTCCAT 659

RESULT 11  
US-08-816-105A-2/c  
Sequence 2, Application US/08816105A  
Patent No. 5989882  
GENERAL INFORMATION:  
APPLICANT: Crueger, Anneliese; Dellweg, Hans-Georg; Lenz,  
APPLICANT: Schroder, Werner; Pape, Hermann;  
APPLICANT: Goeke, Klaus; Schaper, Beate; Hemker, Michael;  
APPLICANT: Piepersberg, Wolfgang; Distler, Jurgen;  
APPLICANT: Stratmann, Ansgar  
TITLE OF INVENTION: PROCESSES FOR PREPARING ACARVIOXYL

;; TITLE OF INVENTION: TRANSFERASE AND FOR USING IT IN THE  
;; TITLE OF INVENTION: CONVERSION OF ACARBOSE HOMOLOGUES  
;; TITLE OF INVENTION: INFO ACARBOSE, FOR THE PREPARATION  
;; TITLE OF INVENTION: OF ACARBOSE HOMOLOGUES  
;; NUMBER OF SEQUENCES: 16  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Sprung Kramer Schaefer & Briscoe  
;; STREET: 660 White Plains Road  
;; CITY: Tarrytown  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10591-5144

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb. storage  
;; COMPUTER: Apple Macintosh 6500  
;; OPERATING SYSTEM: System 7.5  
;; SOFTWARE: Wordperfect 3.5

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/816,105A  
;; FILING DATE: 14-MAR-1997

;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: DE 19625269.5  
;; FILING DATE: 25-JUN-1996

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: DE 19611252.4  
;; FILING DATE: 22-MAR-1996

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kurt G. Briscoe  
;; REGISTRATION NUMBER: 33,141

;; REFERENCE/DOCKET NUMBER: Bayer 9814-KGB  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (914) 332-1700  
;; TELEFAX: (914) 332-1844

;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2582 nucleotides  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear

US-08-816-105A-2

Query Match 77.9%; Score 14.8; DB 2; Length 2582;  
Best Local Similarity 88.9%; Pred. No. 2e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCAT 18  
Db 635 GTTCGCCATCCAGTCCAT 618

RESULT 12  
US-09-031-563-1  
; Sequence 1, Application. US/09031563A  
; Patent No. 6022708  
; GENERAL INFORMATION:  
; APPLICANT: Frederic de Sauvage  
; APPLICANT: Arnon Rosenthal  
; TITLE OF INVENTION: Fused  
; FILE REFERENCE: P1272  
; CURRENT FILING DATE: 1998-02-26  
; CURRENT APPLICATION NUMBER: US/09/031,563A  
; NUMBER OF SEQ ID NOS: 27

SEQ ID NO 1  
; LENGTH: 4880  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unknown N  
; LOCATION: 4160, 4243, 4361  
; OTHER INFORMATION: unknown  
US-09-031-563-1

Query Match 77.9%; Score 14.8; DB 3; Length 4880;  
Best Local Similarity 88.9%; Pred. No. 2.1e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TTGCGCATCCAGTCCATC 19  
Db 2866 TTGCGCATCCAGTCCACC 2883

RESULT 13  
US-09-392-277-1  
; Sequence 1, Application US/09392277A  
; Patent No. 6451977  
; GENERAL INFORMATION:  
; APPLICANT: Frederic de Sauvage  
; APPLICANT: Arnon Rosenthal  
; TITLE OF INVENTION: Fused  
; FILE REFERENCE: P1272R1P1  
; CURRENT APPLICATION NUMBER: US/09/392,277A  
; CURRENT FILING DATE: 1999-09-03  
; EARLIER APPLICATION NUMBER: US 09/258,000  
; EARLIER FILING DATE: 1999-02-25  
; EARLIER APPLICATION NUMBER: US 60/076,072  
; EARLIER FILING DATE: 1998-02-26  
; NUMBER OF SEQ ID NOS: 28  
; SEQ ID NO 1  
; LENGTH: 4880  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unknown N  
; LOCATION: 4160, 4243, 4361  
; OTHER INFORMATION: unknown  
US-09-392-277-1

Query Match 77.9%; Score 14.8; DB 4; Length 4880;  
Best Local Similarity 88.9%; Pred. No. 2.1e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TTGCGCATCCAGTCCATC 19  
Db 2866 TTGCGCATCCAGTCCACC 2883

RESULT 14  
US-09-031-563-4  
; Sequence 4, Application US/09031563A  
; Patent No. 6022708  
; GENERAL INFORMATION:  
; APPLICANT: Frederic de Sauvage  
; APPLICANT: Arnon Rosenthal  
; TITLE OF INVENTION: Fused  
; FILE REFERENCE: P1272  
; CURRENT APPLICATION NUMBER: US/09/031,563A  
; CURRENT FILING DATE: 1998-02-26  
; NUMBER OF SEQ ID NOS: 27  
; SEQ ID NO 4  
; LENGTH: 5125  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-031-563-4

Query Match 77.9%; Score 14.8; DB 3; Length 5125;  
Best Local Similarity 88.9%; Pred. No. 2.1e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TTGCGCATCCAGTCCATC 19  
Db 3110 TTGCGCATCCAGTCCACC 3127

RESULT 15  
US-09-392-277-4

; Sequence 4, Application US/09392277A  
; Patent No. 6451977  
; GENERAL INFORMATION:  
; APPLICANT: Frederic de Sauvage  
; APPLICANT: Arnon Rosenthal  
; TITLE OF INVENTION: Fused  
; FILE REFERENCE: P1272R1P1  
; CURRENT APPLICATION NUMBER: US/09/392,277A  
; CURRENT FILING DATE: 1999-09-03  
; EARLIER APPLICATION NUMBER: US 09/258,000  
; EARLIER FILING DATE: 1999-02-25  
; EARLIER APPLICATION NUMBER: US 60/076,072  
; EARLIER FILING DATE: 1998-02-26  
; NUMBER OF SEQ ID NOS: 28  
; SEQ ID NO 4  
; LENGTH: 5125  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-392-277-4

Query Match 77.9%; Score 14.8; DB 4; Length 5125;  
Best Local Similarity 88.9%; Pred. No. 2.1e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TTCGCCATCCAGTCCATC 19  
|||||  
Db 3110 TTCGCTATCCAGTCCACC 3127

Search completed: July 8, 2003, 09:32:42  
Job time : 27.2793 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 19:09:54 ; Search time 109.737 Seconds  
(without alignments)  
273.390 Million cell updates/sec

Title: US-09-647-780A-20

Perfect score: 19

Sequence: 1 gttccgcatccagtcac 19

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1105431 seqs, 789497651 residues.

Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17.4	91.6	1226	9	US-10-106-698-2099
2	17.4	91.6	2893	9	US-10-017-273A-4
3	17.4	91.6	2893	10	US-09-905-846-1
4	17.4	91.6	2975	9	US-10-017-273A-5
5	17.4	91.6	2975	10	US-09-905-846-5
6	16.4	86.3	1431	9	US-09-908-988B-1
7	16.4	84.2	558	9	US-10-156-761-4649
8	16.4	84.2	9025608	9	US-10-156-761-1
9	15.8	83.2	477	9	US-09-918-995-2051
10	15.8	83.2	548	10	US-09-815-343-386
11	15.8	83.2	642	10	US-09-974-300-2388
12	15.8	83.2	1194	9	US-10-156-761-5444
13	15.8	83.2	1479	9	US-09-844-948-3
14	15.8	83.2	2446	9	US-10-139-662-11
15	15.8	83.2	2446	9	US-10-139-683-11
16	15.8	83.2	2557	9	US-09-844-948-1
17	15.8	83.2	3583	9	US-10-084-817-202
18	15.8	83.2	10115	9	US-09-764-891-7941
19	15.8	83.2	9025608	9	US-10-156-761-1

20	15.4	81.1	2055	9	US-10-121-032-58
21	15.4	81.1	2055	9	US-10-093-037-58
22	15.4	81.1	2975	10	US-09-917-800A-1418
23	15	78.9	152	10	US-09-864-761-17150
24	15	78.9	224	10	US-09-960-352-446
25	15	78.9	445	10	US-09-864-761-337
26	15	78.9	1533	9	US-10-156-761-4106
27	14.8	77.9	330	10	US-09-563-817-900
28	14.8	77.9	387	9	US-10-156-761-7326
29	14.8	77.9	447	9	US-09-918-995-32026
30	14.8	77.9	456	9	US-09-918-995-32021
31	14.8	77.9	466	9	US-09-918-995-12229
32	14.8	77.9	672	10	US-09-770-149-408
33	14.8	77.9	707	10	US-09-910-943-253
34	14.8	77.9	1252	9	US-10-118-495-1
35	14.8	77.9	1413	10	US-09-822-830A-469
36	14.8	77.9	1458	9	US-10-128-714-7011
37	14.8	77.9	1513	9	US-10-128-714-6011
38	14.8	77.9	1941	10	US-09-974-300-963
39	14.8	77.9	2000	9	US-10-198-846-10290
40	14.8	77.9	2209	10	US-09-880-107-3666
41	14.8	77.9	2760	10	US-09-070-927A-413
42	14.8	77.9	3275	9	US-10-128-714-11
43	14.8	77.9	3301	9	US-10-092-154-1718
44	14.8	77.9	3301	10	US-09-764-847-1718
45	14.8	77.9	3513	9	US-10-128-714-5011

#### ALIGNMENTS

##### RESULT 1

US-10-106-698-2099  
; Sequence 2099, Application US/101066698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides  
; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; PRIOR FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 2099  
; LENGTH: 1226  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (164)..(164)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (1180)..(1180)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (1197)..(1197)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (1215)..(1215)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (1221)..(1221)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (1224)..(1224)  
; OTHER INFORMATION: n equals a,t,g, or c  
; US-10-106-698-2099

Sequence 58, Appl  
Sequence 58, Appl  
Sequence 1418, Ap  
Sequence 17150, A  
Sequence 446, App  
Sequence 337, App  
Sequence 4106, Ap  
Sequence 900, App  
Sequence 7326, Ap  
Sequence 32026, A  
Sequence 32021, A  
Sequence 12229, A  
Sequence 408, App  
Sequence 253, App  
Sequence 1, Appli  
Sequence 469, App  
Sequence 7011, Ap  
Sequence 6011, Ap  
Sequence 963, App  
Sequence 10290, A  
Sequence 3666, Ap  
Sequence 413, App  
Sequence 11, Appl  
Sequence 1718, Ap  
Sequence 5011, Ap

Query Match 91.6%; Score 17.4; DB 9; Length 1226;  
Best Local Similarity 94.7%; Pred. No. 18;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19  
|||||  
DB 606 GTTCGCCATCCATCCATC 624

RESULT 2

US-10-017-273A-4  
; Sequence 4, Application US/10017273A  
; Publication No. US20030119714A1  
; GENERAL INFORMATION:  
; APPLICANT: Pfizer Inc.  
; APPLICANT: Naylor, Alasdair M.  
; APPLICANT: Van Der Graaf, Pieter H  
; APPLICANT: Wayman, Christopher P.  
; TITLE OF INVENTION: Treatment of Male Sexual Dysfunction  
; FILE REFERENCE: PC22013  
; CURRENT APPLICATION NUMBER: US/10/017,273A  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/265,358  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: GB 0030647.2  
; PRIOR FILING DATE: 2000-12-15  
; PRIOR APPLICATION NUMBER: GB 0108730.3  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: GB 0120679.6  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 09/905,846  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/291,722  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: US 09/895,367  
; PRIOR FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 2893  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-017-273A-4

Query Match 91.6%; Score 17.4; DB 9; Length 2893;  
Best Local Similarity 94.7%; Pred. No. 18;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19  
|||||  
DB 2374 GTTCGCCATCCATCCATC 2392

RESULT 3

US-09-905-846-1  
; Sequence 1, Application US/09905846  
; Patent No. US20020102707A1  
; GENERAL INFORMATION:  
; APPLICANT: Ian Dennis Harrow  
; APPLICANT: Peter Stacey  
; APPLICANT: Roderick Thomas Walsh  
; TITLE OF INVENTION: Compounds for the treatment of sexual dysfunction  
; FILE REFERENCE: PCS10926APME  
; CURRENT APPLICATION NUMBER: US/09/905,846  
; CURRENT FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 0017387.2  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/220,908  
; PRIOR FILING DATE: 2000-07-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1

LENGTH: 2893  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-905-846-1

Query Match 91.6%; Score 17.4; DB 10; Length 2893;  
Best Local Similarity 94.7%; Pred. No. 18;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19  
|||||  
DB 2374 GTTCGCCATCCATCCATC 2392

RESULT 4

US-10-017-273A-5  
; Sequence 5, Application US/10017273A  
; Publication No. US20030119714A1  
; GENERAL INFORMATION:  
; APPLICANT: Pfizer Inc.  
; APPLICANT: Naylor, Alasdair M.  
; APPLICANT: Van Der Graaf, Pieter H  
; APPLICANT: Wayman, Christopher P.  
; TITLE OF INVENTION: Treatment of Male Sexual Dysfunction  
; FILE REFERENCE: PC22013  
; CURRENT APPLICATION NUMBER: US/10/017,273A  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/265,358  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: GB 0030647.2  
; PRIOR FILING DATE: 2000-12-15  
; PRIOR APPLICATION NUMBER: GB 0108730.3  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: GB 0120679.6  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 09/905,846  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/291,722  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: US 09/895,367  
; PRIOR FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 2975  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-017-273A-5

Query Match 91.6%; Score 17.4; DB 9; Length 2975;  
Best Local Similarity 94.7%; Pred. No. 18;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19  
|||||  
DB 2439 GTTCGCCATCCATCCATC 2457

RESULT 5

US-09-905-846-5  
; Sequence 5, Application US/09905846  
; Patent No. US20020102707A1  
; GENERAL INFORMATION:  
; APPLICANT: Ian Dennis Harrow  
; APPLICANT: Peter Stacey  
; APPLICANT: Roderick Thomas Walsh  
; TITLE OF INVENTION: Compounds for the treatment of sexual dysfunction  
; FILE REFERENCE: PCS10926APME  
; CURRENT APPLICATION NUMBER: US/09/905,846  
; CURRENT FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 0017387.2  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/220,908

; PRIOR FILING DATE: 2000-07-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 2975  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-905-846-5

Query Match 91.6%; Score 17.4; DB 10; Length 2975;  
Best Local Similarity 94.7%; Pred. No. 18;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19  
||| ||||| ||||| ||||| |||||  
DB 2439 GTTCGCCATCCATCCATC 2457

RESULT 6  
US-09-908-988B-1  
; Sequence 1, Application US/09908988B  
; Patent No. US20020127690A1  
; GENERAL INFORMATION:  
; APPLICANT: OLSON, ERIC  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES  
; FILE REFERENCE: IN STRIATED MUSCLE CELLS  
; CURRENT FILING DATE: 2000-07-18  
; CURRENT APPLICATION NUMBER: US/09/908,988B  
; PRIOR FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: 60/219,020  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1431  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (199)..(1296)  
US-09-908-988B-1

Query Match 86.3%; Score 16.4; DB 10; Length 1431;  
Best Local Similarity 94.4%; Pred. No. 56;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCAT 18  
||| ||||| ||||| ||||| |||||  
DB 993 GTTCGCCATCCAGTCCAT 1010

RESULT 7  
US-10-156-761-4649/c  
; Sequence 4649, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 4649  
; LENGTH: 558  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(558)  
US-10-156-761-4649

Query Match 84.2%; Score 16; DB 9; Length 558;  
Best Local Similarity 100.0%; Pred. No. 90;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CGCCATCCAGTCCATC 19  
||||| ||||| ||||| ||||| |||||  
DB 216 CGCCATCCAGTCCATC 201

RESULT 8  
US-10-156-761-1/c  
; Sequence 1, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 1  
; LENGTH: 9025608  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (4187715)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
US-10-156-761-1

Query Match 84.2%; Score 16; DB 9; Length 9025608;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CGCCATCCAGTCCATC 19  
||||| ||||| ||||| ||||| |||||  
DB 5690362 CGCCATCCAGTCCATC 5690347

RESULT 9  
US-09-918-995-2051/c  
; Sequence 2051, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hysq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2051

LENGTH: 477  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(477)  
OTHER INFORMATION: n = A,T,C or G

US-09-918-995-2051

Query Match 83.2%; Score 15.8; DB 9; Length 477;  
Best Local Similarity 89.5%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19

DB 223 GTTCGCCATCCAGTCCCTC 205

RESULT 10

US-09-815-343-386  
Sequence 386, Application US/09815343

Patent No. US2001005596A1

GENERAL INFORMATION:

APPLICANT: Meagher, Madeleine

APPLICANT: Xu, Jiangchun

APPLICANT: King, Gordon E.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND

FILE OF INVENTION: DIAGNOSIS OF COLON CANCER

FILE REFERENCE: 210121.504

CURRENT APPLICATION NUMBER: US/09/815,343

PRIOR FILING DATE: 2001-03-22

NUMBER OF SEQ ID NOS: 1556

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 386

LENGTH: 548

TYPE: DNA

ORGANISM: Homo sapien

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)...(548)

OTHER INFORMATION: n = A,T,C or G

US-09-815-343-386

Query Match 83.2%; Score 15.8; DB 10; Length 548;  
Best Local Similarity 89.5%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19

DB 16 GTTCGCCATCCAGGACATC 34

RESULT 11

US-09-974-300-2388/c

Sequence 2388, Application US/09974300

Patent No. US20020146721A1

GENERAL INFORMATION:

APPLICANT: Berka, Randy M.

APPLICANT: Clausen, Ib Groth

TITLE OF INVENTION: Methods For Monitoring Multiple Gene

FILE REFERENCE: Expression

CURRENT APPLICATION NUMBER: US/09/974,300

PRIOR FILING DATE: 2001-10-05

NUMBER OF SEQ ID NOS: 8481

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2388

LENGTH: 642

TYPE: DNA

ORGANISM: Bacillus licheniformis  
US-09-974-300-2388

Query Match 83.2%; Score 15.8; DB 10; Length 642;  
Best Local Similarity 89.5%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19

DB 429 GATCGCCATCCAGCCCATC 411

RESULT 12

US-10-156-761-5444

Sequence 5444, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 5444

LENGTH: 1194

TYPE: DNA

ORGANISM: Streptomyces avermitilis

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(1194)

US-10-156-761-5444

Query Match 83.2%; Score 15.8; DB 9; Length 1194;  
Best Local Similarity 89.5%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19

DB 1032 GTTCGCCATCCAGTCCCTC 1050

RESULT 13

US-09-844-948-3/c

Sequence 3, Application US/09844948

Publication No. US20030119161A1

GENERAL INFORMATION:

APPLICANT: Meyers, Rachel A.

APPLICANT: Williamson, Mark

TITLE OF INVENTION: 32132, A NOVEL FUCOSYLTRANSFERASE FAMILY

FILE REFERENCE: 10448-048001

CURRENT APPLICATION NUMBER: US/09/844,948

CURRENT FILING DATE: 2001-04-24

PRIOR APPLICATION NUMBER: US 60/200,604

PRIOR FILING DATE: 2000-04-28

NUMBER OF SEQ ID NOS: 6

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

LENGTH: 1479

TYPE: DNA

ORGANISM: Homo sapiens

US-09-844-948-3

Query Match 83.2%; Score 15.8; DB 9; Length 1479;  
Best Local Similarity 89.5%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GTTCGCCATCCAGTCCATC 19  
||||| ||||||| |||  
Db 951 GTTCGCCATCCAGTCCCTC 933

## RESULT 14

US-10-139-662-11  
; Sequence 11, Application US/10139662  
; Publication No. US20030027358A1  
; GENERAL INFORMATION:  
; APPLICANT: Cosman, David J.  
; APPLICANT: Anderson, Dirk M.  
; APPLICANT: Borges, Luis  
; TITLE OF INVENTION: Family of Immunoregulators Designated Leukocyte Immunoglobulin-  
; TITLE OF INVENTION: Like Receptors (LIR)  
; FILE REFERENCE: 2624-A  
; CURRENT APPLICATION NUMBER: US/10/139,662  
; CURRENT FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: 09/310,463  
; PRIOR FILING DATE: 1999-05-12  
; PRIOR APPLICATION NUMBER: 08/842,248  
; PRIOR FILING DATE: 1997-04-24  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 2446  
; TYPE: DNA  
; ORGANISM: human  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (171)..(1040)  
US-10-139-662-11

Query Match 83.2%; Score 15.8; DB 9; Length 2446;  
Best Local Similarity 89.5%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GTTCGCCATCCAGTCCATC 19  
||||| ||||||| |||  
Db 416 GTTCGCCATCCGTCATC 434

## RESULT 15

US-10-139-683-11  
; Sequence 11, Application US/10139683  
; Publication No. US20030060614A1  
; GENERAL INFORMATION:  
; APPLICANT: Cosman, David J.  
; APPLICANT: Anderson, Dirk M.  
; APPLICANT: Borges, Luis  
; TITLE OF INVENTION: Family of Immunoregulators Designated Leukocyte Immunoglobulin-  
; TITLE OF INVENTION: Like Receptors (LIR)  
; FILE REFERENCE: 2624-A  
; CURRENT APPLICATION NUMBER: US/10/139,683  
; CURRENT FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: 09/310,463  
; PRIOR FILING DATE: 1999-05-12  
; PRIOR APPLICATION NUMBER: 08/842,248  
; PRIOR FILING DATE: 1997-04-24  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 2446  
; TYPE: DNA  
; ORGANISM: human  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (171)..(1040)

## US-10-139-683-11

Query Match 83.2%; Score 15.8; DB 9; Length 2446;  
Best Local Similarity 89.5%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GTTCGCCATCCAGTCCATC 19  
||||| ||||||| |||  
Db 416 GTTCGCCATCCGTCATC 434

Search completed: July 9, 2003, 02:22:52  
Job time : 124.787 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2003, 00:47:28 ; Search time 963.254 Seconds  
(without alignments)  
319.453 Million cell updates/sec

Title: US-09-647-780A-20

Perfect score: 19  
Sequences: 1 gttccatccagtcacatc 19

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estmu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_hic:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_hic:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	gb_gss:*
18:	em_gss_hum:*
19:	em_gss_inv:*
20:	em_gss_pin:*
21:	em_gss_vrt:*
22:	em_gss_fun:*
23:	em_gss_mam:*
24:	em_gss_mus:*
25:	em_gss_other:*
26:	em_gss_pro:*
27:	em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	19	100.0	507	BE106100	BE106100 UI-R-B01-
C 2	17.4	91.6	351	12 BF987421	BF987421 MR3-GN015
C 3	17.4	91.6	481	10 AW924814	AW924814 WSL 72 C0
C 4	17.4	91.6	524	10 AW745327	AW745327 WSL 33 G0
C 5	17.4	91.6	530	12 BF446399	BF446399 7p36g07.x
C 6	17.4	91.6	554	12 BF437289	BF437289 7p69f03.x

C 7	17.4	91.6	561	13 BG933273	BG933273 WSL 92_D0
C 8	17.4	91.6	576	10 BE355985	BE355985 DGL 121_C
C 9	17.4	91.6	592	10 BE594778	BE594778 PTL 45_D0
C 10	17.4	91.6	704	12 BG153322	BG153322 nag12d11.
C 11	17	89.5	661	17 BH350313	BH350313 CH230-320
C 12	16.4	86.3	255	9 AA501189	AA501189 vnr75b04.r
C 13	16.4	86.3	291	13 BI009951	BI009951 MR4-RT004
C 14	16.4	86.3	303	10 BE144095	BE144095 MR0-HT016
C 15	16.4	86.3	303	10 BE144204	BE144204 MR0-HT016
C 16	16.4	86.3	311	12 BG842100	BG842100 MEST36-B1
C 17	16.4	86.3	360	12 BF590391	BF590391 naa36a05.
C 18	16.4	86.3	375	9 AL777580	AL777580 AL777580
C 19	16.4	86.3	390	9 AI613878	AI613878 vnr75b04.y
C 20	16.4	86.3	439	14 BM868098	BM868098 mgcs013xg
C 21	16.4	86.3	456	10 BE099923	BE099923 UI-R-BJ1-
C 22	16.4	86.3	461	9 AV006036	AV006036 AV006036
C 23	16.4	86.3	485	11 AK019655	AK019655 Mus muscu
C 24	16.4	86.3	504	9 AI712700	AI712700 UI-R-AFL-
C 25	16.4	86.3	571	17 BH816223	BH816223 AM_BA002
C 26	16.4	86.3	608	12 BF284900	BF284900 EST449491
C 27	16.4	86.3	621	13 BJ183928	BJ183928 BJ183928
C 28	16.4	86.3	638	9 AA800245	AA800245 EST189742
C 29	16.4	86.3	641	10 BB662700	BB662700 BB662700
C 30	16.4	86.3	659	14 BQ929484	BQ929484 AGENCOURT
C 31	16.4	86.3	663	13 BI959826	BI959826 HVSMEH002
C 32	16.4	86.3	885	12 BG443392	BG443392 GA_EA002
C 33	16.4	86.3	927	12 BF179418	BF179418 601806987
C 34	16.4	86.3	943	13 BI102877	BI102877 602888472
C 35	16.4	86.3	957	11 AK016235	AK016235 Mus muscu
C 36	16.4	86.3	991	17 CNS0502F	AL348720 Tetraodon
C 37	16	84.2	207	14 BQ820536	BQ820536 1030085E0
C 38	16	84.2	325	13 BI725875	BI725875 1031081F0
C 39	16	84.2	325	13 BI725876	BI725876 1031081F0
C 40	16	84.2	325	13 BM000030	BM000030 1031081F0
C 41	16	84.2	402	14 BQ821766	BQ821766 1030094H0
C 42	16	84.2	461	10 BE403532	BE403532 WHE0427_F
C 43	16	84.2	494	10 BE444568	BE444568 WHE1126_B
C 44	16	84.2	512	12 BG846660	BG846660 1024014E0
C 45	16	84.2	520	14 BQ819945	BQ819945 1030081C0

#### ALIGNMENTS

RESULT 1  
BE106100/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT

BE106100 507 bp mRNA linear EST 13-JUN-2000  
UI-R-B01-ask-h-09-0-UI.s1 UI-R-B01 Rattus norvegicus cDNA clone  
UI-R-B01-ask-h-09-0-UI 3', mRNA sequence.

BE106100  
BE106100.1 GI:8498202  
EST.  
Norway rat.  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

1 (bases 1 to 507)  
Bonaldi,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477

Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA present in the NotI site and the oligo-dT track served to identify it as a clone from the normalized medulla library cDNA library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)  
Seq primer: M13 Forward.  
POLYA=Yes.

## FEATURES

source

Location/Qualifiers

1..507  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-B01-ask-h-09-0-UI"  
/clone\_lib="UI-R-B01"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73B-pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; The library (UI-R-B01) is a subtracted library derived from a mixture of the following tissues: thalamus, cerebellum, hypothalamus, medulla, pons, midbrain, cerebral cortex, corpus striatum and hippocampus. For a detailed description of the library from which this clone was derived, please visit our web site at [rategeng.uiowa.edu](http://rategeng.uiowa.edu). The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
TAG\_LIB="UI-R-B01"  
TAG\_TISSUE="medulla"  
TAG\_SEQ="GAACCG"

BASE COUNT 105 a 133 c 158 g 111 t  
ORIGIN

Query Match 100.0%; Score 19; DB 10; Length 507;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19  
Db 468 GTTCGCCATCCAGTCCATC 450

## RESULT 2

BF987421/c  
LOCUS BF987421 351 bp mRNA linear EST 23-JAN-2001  
DEFINITION MR3-GN0150-101000-002-c07 GN0150.Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF987421  
VERSION BF987421.1 GI:12393743  
KEYWORDS EST.  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 351)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W.Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., Gdeoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

## TITLE

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663

## JOURNAL

## MEDLINE

## COMMENT

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR3&td=MR3-GN0150-101000-002-c07&ts=2000-10-10&td=1>)  
Seq primer: puc 18 forward  
High quality sequence start: 32  
High quality sequence stop: 351.

## FEATURES

source

Location/Qualifiers

1..351  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="GN0150"  
/dev\_stage="Adult"  
/note="Organ: placenta normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 84 a 99 c 94 g 74 t  
ORIGIN

Query Match 91.6%; Score 17.4; DB 12; Length 351;  
Best Local Similarity 94.7%; Pred. No. 9.6e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19  
Db 86 GTTCGCCATCCAGTCCATC 68

## RESULT 3

AW924814/c

## LOCUS

AW924814 481 bp mRNA linear EST 19-JUL-2000  
DEFINITION WSL\_72\_C08\_b1\_A002 Water-stressed 1 (WS1) Sorghum bicolor cDNA, mRNA sequence.

ACCESSION AW924814

VERSION AW924814.1 GI:8090640

KEYWORDS EST.

SOURCE sorghum.

## ORGANISM

Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.

## REFERENCE

## AUTHORS

1 (bases 1 to 481)  
Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt, L.H.

## TITLE

## JOURNAL

## COMMENT

An EST database from Sorghum: water-stressed plants  
Unpublished (2000)  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: [mmpratt@uga.edu](mailto:mmpratt@uga.edu)  
Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.

Seq primer: JEN REV

High quality sequence stop: 479

POLYA-No.

## FEATURES

source

Location/Qualifiers

1..481  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Water-stressed 1 (WS1)"  
/note="Organ: Mix of 5-week old plants on days 7 & 8 after water was withheld; Vector: Lambda zap; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision."

BASE COUNT

82 a 154 c 172 g 73 t





CDNA Library Preparation: M. Bento Soares, Ph.D.  
 DNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL, send email to:  
 info@image.llnl.gov  
 Seq primer: -400P from Gibco  
 High quality sequence stop: 462.

Location/Qualifiers  
 1. 554  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3651197"  
 /clone\_lib="NCI-CGAP\_Pr28"  
 /sex="male"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)  
 with a modified polylinker; Plasmid DNA from the  
 normalized library NCI-CGAP\_Pr22 was prepared, and as  
 circles were made in vitro. Following HAP purification,  
 this DNA was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from a pool  
 of 5,000 clones made from the same library (cloneIDs  
 985608-986759, 1101192-1101959, and 1217928-1220615).  
 Subtraction by Bento Soares and M. Fatima Bonaldo."

FEATURES  
 source

BASE COUNT 93 a 157 c 194 g 110 t  
 ORIGIN

Query Match 91.6%; Score 17.4; DB 12; Length 554;  
 Best Local Similarity 94.7%; Pred. No. 1.1e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19  
 |||||

Db 504 GTTCGCCATCCATCCATC 486

RESULT 7  
 BG933273/c  
 LOCUS  
 DEFINITION WSI\_92\_D06.g1\_A002 Water-stressed 1 (WS1) Sorghum bicolor cDNA,  
 mRNA sequence.  
 BG933273  
 VERSION BG933273.1 GI:14328405  
 KEYWORDS EST.  
 SOURCE sorghum.

ORGANISM  
 Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE 1 (bases 1 to 561)  
 AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt  
 ,L.H.

TITLE An EST database from Sorghum: water-stressed plants  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions  
 below Phred quality 16. The threshold for highest quality sequence  
 is 20.

Seq primer: PolyTMix  
 High quality sequence start: 31  
 High quality sequence stop: 555  
 POLYA=No.

FEATURES  
 source  
 1. 561  
 /organism="Sorghum bicolor"

/db\_xref="taxon:4558"  
 /clone\_lib="Water-stressed 1 (WS1)"  
 /note="Organ: Mix of 5-week old plants on days 7 & 8 after  
 water was withheld; Vector: Lambda Zap; Site\_1: XhoI;  
 Site\_2: EcoRI; The library was made from poly-A RNA in the  
 cloning vector lambda ZAP II. Clones to be sequenced were  
 prepared by mass excision."

BASE COUNT 119 a 136 c 170 g 136 t  
 ORIGIN

Query Match 91.6%; Score 17.4; DB 13; Length 561;  
 Best Local Similarity 94.7%; Pred. No. 1.1e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19  
 |||||

Db 125 GTTCGCCATCCAGTCCATC 107

RESULT 8  
 BE355985/c  
 LOCUS  
 DEFINITION DGL\_121\_C06.b1\_A002 Dark Grown 1 (DGL) Sorghum bicolor cDNA, mRNA  
 sequence.  
 BE355985  
 VERSION BE355985.1 GI:9297542  
 KEYWORDS EST.  
 SOURCE sorghum.

ORGANISM  
 Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE 1 (bases 1 to 576)  
 AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt  
 ,L.H.

TITLE An EST database from Sorghum: dark-grown seedlings  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions  
 below Phred quality 16. The threshold for highest quality sequence  
 is 20.

Seq primer: JEN REV  
 High quality sequence stop: 511  
 POLYA=No.

FEATURES  
 source  
 1. 576  
 /organism="Sorghum bicolor"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Dark Grown 1 (DGL)"

/note="Organ: 5-day-old dark-grown seedlings; Vector:  
 Lambda Zap; Site\_1: XhoI; Site\_2: EcoRI; The library was  
 made from poly-A RNA in the cloning vector lambda ZAP II.  
 Clones to be sequenced were prepared by mass excision."

BASE COUNT 106 a 178 c 203 g 89 t  
 ORIGIN

Query Match 91.6%; Score 17.4; DB 10; Length 576;  
 Best Local Similarity 94.7%; Pred. No. 1.1e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCGCCATCCAGTCCATC 19  
 |||||

Db 353 GTTCGCCATCCAGTCCATC 335

RESULT 9  
 BE594778/c



Clones are derived from the rat BAC library CHORI-230  
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library  
availability, please contact Pieter de Jong (pdejong@mail.choi.org).  
Clones may be purchased from BACPAC Resources  
([http://www.chori.org/bacpac/orering\\_information.htm](http://www.chori.org/bacpac/orering_information.htm)). BAC end  
plate: [http://www.tigr.org/tdb/bac\\_ends/rat/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html)  
Seq primer: SP6  
Class: BAC ends.

## FEATURES

## source

Location/Qualifiers  
1. .561  
/organism="Rattus norvegicus"  
/strain="BN/SSNHsd/MCW"  
/db\_xref="taxon:10116"  
/clone="CH230-32012"  
/clone\_lib="CHORI-230 Segment 1"  
/sex="Female"  
/cell\_type="Brain"  
/note="Vector: pTARBAC2.1; Site\_1: EcoRI; Site\_2: EcoRI;  
CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by  
Pieter de Jong"

## BASE COUNT

## ORIGIN

Query Match 89.5%; Score 17; DB 17; Length 661;  
Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TCGCCATCCAGTCCATC 19

Db 141 TCGCCATCCAGTCCATC 125

## RESULT 12

## AA501189

## LOCUS

DEFINITION 255 bp mRNA linear EST 01-JUL-1997  
vh75b04.r1 Knowles Solter mouse inner cell mass Mus musculus CDNA  
clone IMAGE:892783 5', mRNA sequence.

## ACCESSION

## AA501189

## VERSION

## AA501189.1

## KEYWORDS

## EST.

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS:

## TITLE

## JOURNAL

## COMMENT

## CONTACT

## UNPUBLISHED (1996)

## WashU-HMI Mouse EST Project

## Washington University School

## 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

## Tel: 314 286 1800

## Fax: 314 286 1810

## Email: mouseest@wustl.edu

## This clone is available royalty-free through LLNL; contact the

IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.

## MGI:520743.

## FEATURES

## source

## Location/Qualifiers

## 1. .255

## /organism="Mus musculus"

## /strain="C57BL/6J x DBA/2J F1"

## /db\_xref="taxon:10090"

## /clone="IMAGE:892783"

## /clone\_lib="Knowles Solter mouse inner cell mass"

## /dev\_stage="embryo (post-implantation)"

## /lab\_host="DH10B"

## /note="Vector: pBluescript SK+; Site\_1: XbaI; Site\_2: XhoI"

; Cloned unidirectionally from mRNA prepared from  
primitive streak embryonic tissue. Primer: Oligo dt.  
cDNAs were cloned into the XbaI/XhoI sites of pBluescript  
SK+ (Stratagene) using commercial linkers (NEB). Average  
insert size: 0.5 kb. "

BASE COUNT 70 a 63 c 53 g 69 t  
ORIGIN

Query Match 86.3%; Score 16.4; DB 9; Length 255;  
Best Local Similarity 94.4%; Pred. No. 2.4e+03;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTCCGATCCAGTCCAT 18

Db 42 GTTCCGATCCAGTCCAT 59

## RESULT 13

## BI009951

## LOCUS

DEFINITION 291 bp mRNA linear EST 13-JUN-2001  
MR4-RT0045-250401-008-c05 RT0045 Homo sapiens cDNA, mRNA sequence.

## ACCESSION

## BI009951

## VERSION

## BI009951.1

## KEYWORDS

## EST.

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## COMMENT

## CONTACT

## UNPUBLISHED (2000)

## Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

## Laboratory of Cancer Genetics

## Ludwig Institute for Cancer Research

## Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

## Brazil

## Tel: +55-11-2704922

## Fax: +55-11-2707001

Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)

## This sequence was derived from the FAPESP/LICR Human Cancer Genome

## Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR4&t2=MR4-RT0045-250401-008-c05&t3=2001-04-25&t4=1>)

## Seq primer: puc 18 forward

## High quality sequence start: 73

## High quality sequence stop: 291.

## FEATURES

## source

## Location/Qualifiers

## 1. .291

## /organism="Homo sapiens"

## /db\_xref="taxon:9606"

## /clone\_lib="RT0045"

## /dev\_stage="Adult"

## /note="Organ: kidney tumor; Vector: puc18; Site\_1: SmaI;

## Site\_2: SmaI; A mini-library was made by cloning products

## derived from ORESTES PCR (U.S. Letters Patent application

## No. 196,716 - Ludwig Institute for Cancer Research)

## profiles into the pUC 18 vector. Reverse transcription of

## tissue mRNA and cDNA amplification were performed under

## low stringency conditions."

## BASE COUNT 53 a 128 c 24 g 86 t

## ORIGIN

## Query Match 86.3%; Score 16.4; DB 13; Length 291;

## Best Local Similarity 94.4%; Pred. No. 2.5e+03;

## Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      2 TTGCGCATCCAGTCCATC 19
      ||| ||||| ||||| |||||
Db      202 TTCCCATCCAGTCCATC 219.

RESULT 14
BE144095
LOCUS   BE144095
DEFINITION MRO-HT0165-030200-007-cl2 HT0165 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE144095
VERSION   BE144095.1 GI:8606819
KEYWORDS EST.
SOURCE   human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 303)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
          Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
          Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
          Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
          M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
          Simpson,A.J.
          Shotgun sequencing of the human transcriptome with ORF expressed
          sequence tags
          Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
CONTACT: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-MR0-HT0165-030
          200-007-cl2&t3=2000-02-03&t4=1)
          Seq primer: puc 18 forward
          High quality sequence start: 21
          High quality sequence stop: 182.
          Location/Qualifiers
            1..303
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="HT0165"
              /dev_stage="Adult"
              /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
              Site_2: SmaI; A mini-library was made by cloning products
              derived from ORESTES PCR (U.S. Letters Patent application
              No. 196,716 - Ludwig Institute for Cancer Research)
              profiles into the puc 18 vector. Reverse transcription of
              tissue mRNA and cDNA amplification were performed under
              low stringency conditions."
BASE COUNT 73 a 61 c 57 g 101 t 11 others
ORIGIN

Query Match 86.3%; Score 16.4; DB 10; Length 303;
Best Local Similarity 94.4%; Pred. No. 2.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 TTGCGCATCCAGTCCATC 19
      ||| ||||| ||||| |||||
Db      50 TTCTCCATCCAGTCCATC 67

RESULT 15
BE144204
LOCUS   BE144204
DEFINITION MRO-HT0165-060200-007-cl2 HT0165 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE144204

```

```

VERSION BE144204.1 GI:8606928
KEYWORDS EST.
SOURCE   human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 303)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
          Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
          Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
          Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
          M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
          Simpson,A.J.
          Shotgun sequencing of the human transcriptome with ORF expressed
          sequence tags
          Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
CONTACT: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-MR0-HT0165-060
          200-007-cl2&t3=2000-02-06&t4=1)
          Seq primer: puc 18 forward
          High quality sequence start: 21
          High quality sequence stop: 182.
          Location/Qualifiers
            1..303
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="HT0165"
              /dev_stage="Adult"
              /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
              Site_2: SmaI; A mini-library was made by cloning products
              derived from ORESTES PCR (U.S. Letters Patent application
              No. 196,716 - Ludwig Institute for Cancer Research)
              profiles into the puc 18 vector. Reverse transcription of
              tissue mRNA and cDNA amplification were performed under
              low stringency conditions."
BASE COUNT 73 a 61 c 57 g 101 t 11 others
ORIGIN

Query Match 86.3%; Score 16.4; DB 10; Length 303;
Best Local Similarity 94.4%; Pred. No. 2.6e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 TTGCGCATCCAGTCCATC 19
      ||| ||||| ||||| |||||
Db      50 TTCTCCATCCAGTCCATC 67

Search completed: July 8, 2003, 09:22:17
Job time : 967.404 secs

```